Cell-to-cell and genome-to-genome variability of adenovirus transcription tuned by the cell cycle
Maarit Suomalainen, Vibhu Prasad, Abhilash Kannan and Urs F. Greber
DOI: 10.1242/jcs.252544

Editor: Derek Walsh

Review timeline
Submission to Review Commons: 5 March 2020
Original submission: 3 August 2020
Accepted: 27 August 2020

Reviewer 1

Evidence, reproducibility and clarity

The manuscript by Suomalainen et al. describes a fluorescence-based approach combined reproducibility with high-resolution confocal microscopy to study the heterogeneity of adenovirus infection and clarity in a population of human cells. The main focus of the authors is the detection of viral (Required) transcripts in infected cells, how this correlates with viral genomes, the cell state, and how it varies between different cells in a single population. The paper is generally well written and easy to read, with a few typos, although I found parts of it to be somewhat length and repetitive. Particularly the results section could be pruned somewhat for readability and clarity. The major limitation of the study as it stands is it's overall impact and novelty, which limits journal selection somewhat. A very similar study was recently published, which the authors cite (Krzywkowski et al, 2017). Nevertheless, I think the study design is rigorous and well executed, but I do have some specific comments which may enhance it's overall impact and novelty.

Major:

Results "Visualization of AdV-C5..." section:

Why not also look at normal cells that can be synchronized? Cancer cells, such as A549 will by definition be highly heterogenous and at all phases of the cell cycle. Primary non- transformed cells can easily be synchronized by contact inhibition and are much more physiologically relevant. "The virus particles bound..." - Can the spatial resolution of a confocal microscope truly differentiate individual particles that are sub-wavelength in size? What about the sensitivity for single particles? Some sort of experiment to show that single particles can be detected should be performed and shown to assure the readers that this is in fact possible. Furthermore, even when based on the particle to pfu ratio, the MOI would still be nearly 2000pfu/cell, so the actual number of observed particles is an order of magnitude lower than what was applied to the cells.

Fig. 4 - I am not certain that the observed difference is significant, at least looking at it, beyond the width difference of the peaks, highest expression for both is largely in G1. It would be nice to see this using a western blot of cell cycle sorted cells, which can easily be accomplished using FACS.

Page 15, 2nd paragraph. It would be valuable and informative to determine whether there is heterogeneity in histone association with these different vDNAs and whether these histones exhibit divergent modifications (enabling or restricting transcription).
Same as above. I am rather surprised that the DBP signal did not correlate well with vDNA signal, particularly for the larger replication centers. How can this be reconciled? Was there an increase in overall vDNA signal later in infection? It is important to know this as it determines whether the observed vDNA signal is real or could be caused by viral RNA or other background causes (non-infected controls notwithstanding). Can the signal be detected with inactivated viruses (via UV for example?)

Page 18, 1st paragraph. It would be interesting to determine whether there was association between pol II and those genomes that showed no E1A, similarly to the histone suggestion. What about things like viral chromatin organization? Soriano et al. 2019 showed how E1A and E4orf3 work in tandem to alter viral chromatin organization by varying histone loading on the viral genome.

Fig. 2. Can you really say that a single dot correlates with a single transcript? Has that been validated in any way?

Minor:

Page 5, last paragraph. "Transcripts from the viral late transcription unit,..." This is not correct as recently shown by Crisostomo et al, 2019.

Page 10, "... because AdvV-infected cells are less well adherent..." This is not strictly true as loss of attachment only occurs later on in infection. It would be helpful to have statistical significance indicated directly in the figures.

The very high MOIs used are concerning, could these have negative effects on the cell viability or overall state?

There are a few typos and such that should be corrected

Significance

As I stated above, the work is interesting and significant, to a degree. The major limitation is that the novelty is low as a paper published in 2017 (cited by the authors) used a very similar approach to investigate a similar problem. In addition, there are multiple other recent papers looking at cell populations in the context of adenovirus infection, and whether a single cell or population based approach is better is unclear. This is something the authors might want to strengthen prior to submission.

Reviewer 2

Evidence, reproducibility and clarity

The authors show heterogeneity of AdV-C5 mRNA transcript quantity and dynamics in different cell types, which is regulated by the cell cycle phase and does not correlate to incoming viral DNA, using single molecule RNA FISH technologies and detection of incoming viral DNA by EdC labeling.

Major Comments:

The authors change the MOI used in their experiments (7 different MOIs are used throughout the paper) in a manner that appears randomly and without explanation. (54400 for Figure 1A, 1B, 3B, S3B; 37500 for Figure 1C; 23440 for Figure 2A, 2C, S5A; 13600 for Figure 1A, 1D; 36250 for Figure 3C, S3D; 11200 for Figure 4B; 23400 for Figure 6B). The authors should provide explanation, why these changes in MOIs are necessary.

The authors use mean fluorescence intensity of E1A probes per cell as estimate for viral transcript abundance for some of their experiments (Figure 1D, E, 3B), and count E1A punctae as measure for E1A transcripts in other experiments (Figure 2C, 3C, 5), without showing data, that these measures correlate. Problematic is hereby, that not all E1A punctae have the same signal intensity, as can be seen in Figure S1, which makes the estimation of the correlation of E1A punctae (= number of transcripts) and fluorescence intensity difficult.
The authors should provide both (E1A punctae counts and estimation via fluorescence intensity) for at least one experiment, to prove, that the estimation of E1A transcript levels via fluorescence intensity is feasible.

p.15: “The nuclear E1A signals in AraC-treated cells were resistant to RNase A, but they were dampened by treatment with S1 nuclease (S6B Fig).”

The authors make this statement based on (i) two completely different timepoints (12 h.p.i. for RNaseA treatment, 24.5 h.p.i. for S1 nuclease treatment) and (ii) in different clones of the A549 cells as stated in the methods section on p.21 (Two different clones of human lung epithelial carcinoma A549 cells were used in the study: our laboratory’s old A549 clone (experiments shown in Fig. 1, Fig. 3B and S1 Fig., S3B and S3C Fig., S6A and S6B Fig., RNase A treatment) and A549 from American Type Culture Collection (ATCC, experiments shown in Fig. 2 and Fig. 5, Fig. 6, S2B Fig., S4 Fig., S5 Fig., and S6B Fig. S1 nuclease-treatment)). This makes it difficult to interpret, if the data is due to differences in the timepoints or cell types, or if it is due to binding of the E1A probe to single stranded vDNA.

Minor Comments:

p.4: "AdV are non-enveloped, double-stranded DNA viruses that cause mild respiratory infections in immuno-competent hosts, and establish persistent infections, which can develop into life-threatening infections if the host becomes immuno-compromised [reviewed in 6]."

Not all AdV cause respiratory diseases, the disease outcome of human AdV depends on the site of primary infection, which differs between the different AdV types.

p.7: The authors state, that "At the 17 h time point, about half of the cells had high numbers of protein VI transcripts, and most of them very high numbers of E1A transcripts.”, however, the picture shown in Figure 1F shows a different phenotype, with low transcript levels of VI in E1A high cells and high transcript levels of VI in E1A low cells.

p.8: “This nuclear E1A signal is due to binding of the E1A probe to single-stranded vDNA in the replication centers (see below).”

The authors should state here, that due to the binding of the probes to the single stranded vDNA in the replication centers, the nucleus was excluded from the analysis for Figure 1F in late timepoints. Due to this time point the author cannot state that the E1A staining seen (Fig. 1F; indicated with white arrows) are replication centers; this is just an assumption, since there is no evidence in Fig 1 the author cannot be sure; the author should change the text: “taking the following experiments into account...”, “due to further studies (see below)..... we assume that...”

p.8: The authors should mention the figure they refer to, since there is no E1B-55K staining in Fig. 1F

p.9: Which test was used to calculate the additional p-values?

p.10: For the experiment for the correlation of viral genomes per cell and E1A transcripts in HDF-TERT cells (Figure S2C), the MOI is missing in the description of the results, as well as in the corresponding figure legends.

p. 11: calculation of correlation? rs?

Why does the author combine S and G2/M phase? Fig. S3A show different values for the phases

p.11: “Thus, the total intensity of nuclear DAPI signal can be used to accurately assign G1 vs S/G2/M stage to cells.”

The authors should also here refer to other papers, which showed that this correlation is feasible, as they did in the methods section (67. Roukos V, Pegoraro G, Voss TC, Misteli T. Cell cycle staging of individual cells by fluorescence microscopy. Nature protocols. 2015;10(2):334-48. Epub 2015/01/31. doi: 10.1038/nprot.2015.016. PubMed PMID: 25633629; PubMed Central PMCID:PMC6318798.), and maybe also refer to a newer paper which deals with this technique: Ferro, A., Mestre, T., Carneiro, P. et al. Blue intensity matters for cell cycle profiling in fluorescence DAPI-stained images. Lab Invest 97, 615-625 (2017). https://doi.org/10.1038/labinvest.2017.13
p.11: "Furthermore, when focusing on the highest E1A expressing cells, i.e., the cells with mean cytoplasmic E1A intensities larger than 1.5 × interquartile range from the 75th percentile, 71.9% of these cells were found to be in the G1 phase of cell cycle, whereas only 55.8% of cells in the total sampled cell population were G1 cells."

The authors do not provide any reference to a figure within the manuscript or the supplements, which contains these data. Are these data not shown in the manuscript?

p.12: punctuation mistake; . instead of ,
To enrich G1 cells. AdV-C-5 (moi ~ 36250) was added. Why does the author switch between signal intensities and counting E1A puncta per cell (limited to 200) in the different experiments to illustrate accumulation of E1A transcripts?

p.14: "For E1A (or E1B-55K), we did not detect transcriptional bursts with bDNA-FISH probes on nuclear vDNAs, either prior to or after accumulation of viral transcripts in the cell cytoplasm."

The authors do not provide any reference to a figure within the manuscript or the supplements, which contains these data. Are these data not shown in the manuscript?

p.14: space between number and %

p.15: "This is also seen in AdV-C5-EdC-infected cells" should be changed to "This was also seen in AdV-C5-EdC-infected cells"

Fig. 1B:
-figure legend does not indicate how cells were staine
-also no description in the continuous text

-which E1A transcripts are stained? all? 12S? 13S? Fig. 1D:
-difference in accumulation of viral transcripts is not that visible as in IF staining (Fig. 1B; Fig. 1S);

-graph does not show any difference between E1A and E1B-55K Fig. 1F:
-figure legend does not fit with labelling of IF images and continuous text

-description says 22 h, while IF labeling and text (p. 7, last lane) mentions 23 h pi Fig. 2A:
-figure legend: lane 5

Punctuation wrong: azide-Alexa Fluor488. Alexa Fluor647 Fig. 4A:
-difficulties to understand

-author stated that promoter-driven EGFP expression is clearly dominated by G1 cells for E1A and by S/G2/M cells for CMV, however this is not clearly visible in the graph

-no severe differences visible between CMV-eGFP and E1A-eGFP

-author should include numbers for quantification and statistical calculations to illustrate the differences

Fig. 4B:
-amount of E1A protein levels calculated via IF (signal intensities)

-immunofluorescence is not a suitable tool for protein quantification Fig. 5:
-in A. it is stated, that E1A bDNA -FISH is not suitable, since it is too short to be detectable. However, in B E1A bDNA-FISH is used. Is there a difference?

-according to the method part just one E1A mRNA was used for the assays, why is it then not possible to use that one in Fig. 5A?

-explanation of the procedure and the experiment is very confusing Fig. S6B:
-authors want to show that it is RNase-insensitive, but S1 nuclease-sensitive

two different A549 cell clones and two different time points are used for the treatments? not compareable to each other

Material and Methods:

-headings do not indicate which methods are explained

-no clear structure

Significance

highly significant manuscript
very important for the virology field

my research topics are human adenoviruses and their replication cycle

Reviewer 3

Evidence, reproducibility and clarity

Summary:
Soumalainen et al have studied adenovirus viral gene expression and replication at a single-cell level. They explore the extent of correlation between incoming genome copy number and early gene expression and progression into the late phase, revealing substantial variation between cells in the numbers of E1A transcripts (the first gene expressed upon infection) that is not explained by differences in the numbers of viral genome templates in the cells. They also explore the relevance of cell cycle stage to this variability and show a positive correlation between G1 cell cycle stage and higher levels of gene activity, which explains at least part of the variation.

To form these conclusions they have applied new methods to visualise and quantify single molecules of nucleic acid in single cells. The experiments are all carefully and fully described with full detail of materials. Overall the manuscript is well written and easy to follow.

Major comments:

All of the experiments appear to be done with rigour and their results reported with due regard to statistical significance etc.

My major concern though is that they have been done, perhaps out of necessity to get detectable signals, at very high multiplicities of infection. A well-accepted standard to achieve infection of all cells in a culture is an MOI of 10 infectious units per cell. Even this is acknowledged not to represent the biology of natural infection and it is striking that, where technically feasible, lower MOI studies are more revealing of how a virus actually works.

Here, the authors have used counts of particles rather than infectious units to determine MOI and for Ad5, the particle/pfu ratio is typically 20-100. Their MOIs though are 13,000 - 50,000 per cell, implying an infectious MOI of at least 130 for their A549 experiments, which are known to be readily infected by Ad5 from other work.

Surprisingly, the authors do not see intracellular vDNA copy numbers that are fully reflective of this high MOI, with median intracellular vDNA of 75 /cell at the highest MOI. The authors should consider how the population distribution of vDNA /cell does or does not fit the predicted Poisson distribution. Nonetheless, at these high copy numbers / cell, there must surely be a risk that the variation in gene expression activity arises stochastically, out of competition between genomes for essential transcription factors. Given that multiple cellular factors are each required for E1A transcription, high genome copy numbers could actually inhibit E1A expression relative to cells with more modest copy numbers because limited supplies of individual factors are recruited to different viral genome copies.
It is important for the analysis of correlation of gene expression with cell cycle that the virus has not, at the time point analysed, already perturbed the cell cycle (a well-known effect of infection) which the authors document in Suppl Fig 3B. To my eye, the G1 peak in infected cells is somewhat narrower than in the control while the S/G2 bump is a little greater. The % of cells in each of the two gates needs to be shown to support the conclusion.

Turning to the experiments documenting a correlation between E1A expression and cell cycle stage, the authors interpret their findings in terms of the stage the cells are at when the analysis was done (G1 stage cells have more E1A transcripts). The key experiment (Fig 3B) is analysed at only 4 h pi, so substantial progression from G2/M back to G1 after virus addition can probably be discounted, but the point should be discussed. The authors also use release from G1 in another cell line to support their argument that G1 supports higher levels of E1A expression (Fig 3C). Here, they elect to exclude all cells with fewer than 50 E1A transcripts from their analysis. The reason for this is completely obscure and isn’t obviously justified; conceivably it could bias the outcome of the experiment. At minimum, this decision needs to be carefully explained; ideally, the full data set should be used.

The authors note the highest level of E1A activity (as opposed to RNA) was in G1/S cells and suggest that high E1A cells advance preferentially into S. Whilst in line with the literature that E1A promotes progression into S, an alternative explanation is simply that there is a time lag between RNA accumulation and protein accumulation, during which progression through the cycle would be expected.

Minor comments:

Fig 1 and elsewhere. Given that the 1 h incubations with virus were done at 37 C, the convention would be to include this period in the time post-infection at which harvest / fix time points are quoted. There is inconsistency between text and legend with 12 h pi being sometimes represented as 11 h after virus removal; this is an unnecessary confusion.

Results description prior to the ref to Fig 1B: unclear what this is supposed to mean.

Fig 4A: provide % of cells in each gate in each histogram.

Fig 5: bottom right panel x axis label is wrong

In the presentation of Fig 6, it would be much clearer for the reader if the detected replication foci (ss DNA detected as E1A puncta) were referred to as something other than E1A puncta. There is too much scope for confusion with the earlier experiments in which E1A RNA was detected.

Significance

The study represents the application of state of the art single-molecule visualization techniques to an as yet not understood aspect of virus infection. That said, there is prior experimentation in this area, which the authors fully acknowledge and build upon. The new work is largely descriptive, in that it reveals very clearly the discrepancy between genome copy number and amounts of mRNA without seeking to explain these, beyond the cell cycle analysis. Whilst there is a better correlation between vDNA number and transcript once the data are stratified by cell cycle stage, it is still not strong (Fig 5), indicating that other substantial contributing factors remain to be described.

The work will be of interest certainly to adenovirologists, but also to others who study virus infections - particularly nuclear-replicating DNA viruses such as herpesviruses - where similar considerations are likely to apply.

Expertise: adenovirus; gene expression; virus-host interactions; molecular biology
Author response to reviewers’ comments

The authors would like to thank the reviewers for their comments to the manuscript, and the editor for patience with our response. Our response was delayed due to the COVID-19 lockdown situation in our institution. Now we are pleased to provide the following point-by-point response, as detailed below.

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

The manuscript by Suomalainen et al. describes a fluorescence-based approach combined with high-resolution confocal microscopy to study the heterogeneity of adenovirus infection in a population of human cells. The main focus of the authors is the detection of viral transcripts in infected cells, how this correlates with viral genomes, the cell state, and how it varies between different cells in a single population. The paper is generally well written and easy to read, with a few typos, although I found parts of it to be somewhat lengthy and repetitive. Particularly the results section could be pruned somewhat for readability and clarity. The major limitation of the study as it stands is its overall impact and novelty, which limits journal selection somewhat. A very similar study was recently published, which the authors cite (Krzywkowski et al, 2017). Nevertheless, I think the study design is rigorous and well executed, but I do have some specific comments which may enhance its overall impact and novelty.

Major:

Results “Visualization of AdV-C5...” section:

Why not also look at normal cells that can be synchronized? Cancer cells, such as A549 will by definition be highly heterogenous and at all phases of the cell cycle. Primary non-transformed cells can easily be synchronized by contact inhibition and are much more physiologically relevant.

AU: In the current manuscript, we concentrated on the early phases of the AdV-C5 infection, on the question how virus gene expression is initiated and whether the cell cycle phase of the host cell impacts the initiation of virus gene expression. Answering these questions requires use of cells that express good amount of virus receptors so that viruses efficiently bind to the cells and infections can be synchronized so that extended time does not elapse between virus addition and accumulation of E1A transcripts; extended time between these two steps would make interpretation of the results more complex since cells could have progressed from one cell cycle stage to another during the experiment. Furthermore, having cells at all phases of the cell cycle is actually a benefit since then the experiment can be carried out under an “unperturbed” condition; all cell cycle synchronization methods have pleiotropic effects on the cells.

It is true that primary non-transformed cells are physiologically more relevant than cancer cells, but primary cells have issues with donor-to-donor variability and many primary cells express rather low amounts of AdV-C5 receptors, so synchronized infections in these cells are not possible. Furthermore, the extended cell morphology of many normal fibroblast cell lines and the tendency of cell extensions from neighboring cells to overlap makes fluorescent images of these cells incompatible for automated cell segmentation.

Here, we provide data also from HDF-TERT cells (nontransformed human diploid fibroblasts immortalized by human telomerase expression) to show that two of our key findings from A549 cells are not artefacts of cancer cells. This is, that akin to A549 cells, the infected HDF-TERT cells accumulate high number of E1A transcripts (Fig.1C), and also in these cells nuclear vDNA numbers do not predict the cytoplasmic E1A transcript counts during early phases of infection (S2C Fig). However, since HDF-TERT cells are rather inefficiently infected by AdV-C5, correlation of early E1A transcript accumulation to the cell cycle phase of the host cell could not been done in these cells. We have been unable to identify primary or normal immortalized cells that would be easily available and efficiently infected by AdV-C5 (synchronized infection with short time elapsed between virus addition and accumulation of E1A transcripts).

"The virus particles bound..." - Can the spatial resolution of a confocal microscope truly
differentiate individual particles that are sub-wavelength in size? What about the sensitivity for single particles? Some sort of experiment to show that single particles can be detected should be performed and shown to assure the readers that this is in fact possible. Furthermore, even when based on the particle to pfu ratio, the MOI would still be nearly 2000pfu/cell, so the actual number of observed particles is an order of magnitude lower than what was applied to the cells. AU: The fluorescence signal from individual fluorophore-tagged AdV or anti-hexon antibody-decorated particle is bright enough to be picked up by PMT or HyD detectors of the current confocal laser scanning microscopes. In fact, tracking fluorophore-tagged particles of the size of AdV has been a standard microscopy procedure since late 1990’s.

Because the Reviewers were questioning the apparently high multiplicity of infection used in the experiments, we clarify the difference between “standard” MOI estimations and our infection set-up. First of all, as described in Material and Methods, we estimated the number of physical virus particles in our virus preparations using A260 measurements (J.A. Sweeney et al., Virol. 2002, doi: 10.1006/viro.2002.1406). This method, like all other methods used to estimate virus particle numbers, is likely not 100% reliable.

Second, we incubated the virus inoculum with cells only for 60 min, after which the unbound viruses were washed away. During this short incubation time only a small fraction of input virus particles bind to cells, and indeed as shown in Fig.1A, a theoretical MOI of 54400 physical virus particles/cell or 13600 physical virus particles/cell yielded Median of 75 and 26 bound virus particles per cell, respectively. Interpretation of the results from the cell cycle assays required that there was a relatively short time between infection and analysis so that cells in a large scale did not change their cell cycle status during the experiment. This required use of a rather high MOI. Furthermore, for collection of a large data set, it is convenient that every cell is infected.

Third, what exactly does one pfu mean in terms of physical adenovirus particles? There is no clear answer to this, since several parameters affect the pfu. In which cells was the titration carried out? How long was the input virus inoculum incubated with the cells? How many of the virus particles entering the cell actually established an infection? And, as described in A. Yakimovich et al. (J. Virol. 2012, DOI: 10.1128/JVI.01102-12), only a fraction of infected cells produce a plaque. The majority of papers stating that x pfu/cell was used for infection, usually incubate the cells with the virus inoculum for several hours at 37°C, and never make any attempts to estimate exactly how many virus particles entered into the cells.

Fig. 4 - I am not certain that the observed difference is significant, at least looking at it, beyond the width difference of the peaks, highest expression for both is largely in G1. It would be nice to see this using a western blot of cell cycle sorted cells, which can easily be accomplished using FACS.

AU: In the highest GFP expression bin, CMV-eGFP expressing cells have 43% cells in G1 and 50% in S/G2/M. In comparison, E1A-GFP expressing cells have 58% cells in G1 and 35% in S/G2/M. The difference in G1 cells in the highest egFP bin is statistically significant (p<0.0001; Two Proportion Z-test). This quantification has been added to the main text, as well as to the figure. The calculated p value using a R script was 3.29e-23. We do not think that this needs to be clarified further with FACS sorted cells.

Page 15, 2nd paragraph. It would be valuable and informative to determine whether there is heterogeneity in histone association with these different vDNAs and whether these histones exhibit divergent modifications (enabling or restricting transcription). Same as above, I am rather surprised that the DBP signal did not correlate well with vDNA signal, particularly for the larger replication centers. How can this be reconciled? Was there an increase in overall vDNA signal later in infection? It is important to know this as it determines whether the observed vDNA signal is real or could be caused by viral RNA or other background causes (non-infected controls notwithstanding). Can the signal be detected with inactivated viruses (via UV for example?) AU: Whether histone modifications impact the transcriptional output of adenovirus genomes early in infection is indeed an intriguing question, but unfortunately this is very challenging, if not impossible, to study at single-cell / single vDNA level with the existing technology. Techniques for single-cell measurements of chromatin states are still in infancy, although some notable advancements in this field were reported in 2019 (e.g. K. Grosselin et al. Nature Genetics, DOI: https://doi.org/10.1038/s41588-019-0424-9 and S. Ai et al. Nature Cell Biology,
Furthermore, current literature offers a confused picture as to when exactly protein VII on incoming virus genomes is replaced by histones (reviewed in the reference 39, Giberson et al.). Of note, the vast majority of incoming nuclear vDNA molecules scored protein VII-positive with anti-VII staining under the experimental conditions used for the Fig. 2C data. However, we did not include these results into the manuscript because VII-positive signal on vDNAs does not exclude these vDNAs having histones on certain parts of the genome.

The Reviewer wonders why the DBP signal in Fig. 6C does not correlate with vDNA signal. There is no discrepancy here because DBP signal in the figure is a proxy for replicating vDNA whereas the click vDNA signal reports incoming vDNA. The one DBP spot without an associated click vDNA signal could be due to a replication center originated from a replicated viral genome, not from incoming viral genome. The figure shows that incoming vDNAs within the same nucleus initiate replication asynchronously.

Page 18, 1st paragraph. It would be interesting to determine whether there was association between pol II and those genomes that showed no E1A, similarly to the histone suggestion. What about things like viral chromatin organization? Soriano et al. 2019 showed how E1A and E4orf3 work in tandem to alter viral chromatin organization by varying histone loading on the viral genome.

AU: This again would be technically very challenging to show. We actually tried to visualize active transcription using an antibody against RNA polymerase II CTD repeat YSPTSPS (phosphor S5), azide-alexa fluor488 and anti-alexa fluor488 antibody to mark EdC-labeled incoming vDNAs and proximity ligation assay for signal amplification. However, this method was not sensitive enough to detect RNA polymerase II association with individual viral genomes. We only detected the proximity ligation signal in replication centers when replicated viral genomes were tagged with EdC.

Fig. 2. Can you really say that a single dot correlates with a single transcript? Has that been validated in any way?

AU: Signal amplification with branched DNA technology leads to binding of a large number of fluorescent probes to a mRNA and thus enables detection of single nucleic acid molecules. This has been validated e.g. in A.N. Player et al. 2001. J. Histochem. Cytochem (https://doi.org/10.1177/002215540104900507) and N. Battich et al. 2013. Nature Methods (https://doi.org/10.1038/nmeth.2657).

Minor:

Page 5, last paragraph. "Transcripts from the viral late transcription unit,..." This is not correct as recently shown by Crisostomo et al. 2019.

AU: The data in Crisostomo et al. paper suggest that some late gene expression can occur before vDNA replication, but an abundant accumulation of late transcripts coincides with onset of vDNA replication. However, the Crisostomo et al. study did not test what the levels of late gene transcripts are if the vDNA replication was inhibited. But to acknowledge the possibility that there might be some level of late gene transcription prior to replication of the viral genomes, the sentence is modified as follows:

“Transcripts from the viral late transcription unit, amongst them mRNAs for the viral structural proteins, vastly increase in abundance concomitant with the onset of vDNA replication”.

Furthermore, we have added the Crisostomo et al. reference here as well.

Page 10, "... because AdvV-infected cells are less well adherent..." This is not strictly true as loss of attachment only occurs later on in infection. It would be helpful to have statistical significance indicated directly in the figures.

AU: Although clearly visible cell rounding indeed occurs only late in infection, also during early stages of infection the HADV-C5-infected cells are less adherent than non-infected cells. In many assays this is not obvious, but the RNA FISH staining procedure includes several incubation and washing steps in rather harsh buffers, and we observed random, sometimes considerable, cell loss with infected cultures but not with non-infected cultures.
In the revised manuscript we have included the statistical significance P values both into the main text and the figure legends, but not to the figures directly, because the P values were generated with different statistical tests and P values should not be shown/mentioned without stating which statistical test was used. However, we noticed that we had in some cases omitted to mention what was the number of pairs analyzed in some of the Spearman’s correlation tests. This has now been corrected in the revised manuscript.

The very high MOIs used are concerning, could these have negative effects on the cell viability or overall state?

AU: We refer to our explanation above about the theoretical MOI and the actual MOI. Furthermore, in the experiment described in Fig.2C (correlation of E1A transcripts per cell vs. viral genomes per cell), 42% of analyzed cells had ≤ 5 viral genomes/cell and 27.5% of analyzed cells had between 6-10 viral genomes per cell; these are not high numbers. We also provide controls that the EdC-labeled genomes are detected with good efficiency. Hence the EdC-labeled genomes per cell are a good estimate of the numbers of virus particles that indeed entered into the cells.

There are a few typos and such that should be corrected. AU: We have tried to find and correct the typos.

Reviewer #1 (Significance (Required)):

As I stated above, the work is interesting and significant, to a degree. The major limitation is that the novelty is low as a paper published in 2017 (cited by the authors) used a very similar approach to investigate a similar problem. In addition, there are multiple other recent papers looking at cell populations in the context of adenovirus infection, and whether a single cell or population based approach is better is unclear. This is something the authors might want to strengthen prior to submission.

AU: In the current study, we focused on the early phase of HAdV-C5 infection, on how viral gene expression is initiated and how individual nuclear viral genomes proceed to a replicative phase. The Krzywkowski et al. 2017 J. Virol. Paper that the reviewer refers to used padlock probe-based rolling circle amplification technique to simultaneously detect HAdV-C5 genomes and viral mRNAs in individual infected cells.

The shortcoming of this method is inferior sensitivity compared to the branched DNA technology-based method used by us in the current study. Krzywkowski et al. were able to pick up signals from virus mRNAs and virus genome only relatively late in the infection, i.e. at the time when incoming genomes were expected to have multiplied by replication. Thus the study by Krzywkowski et al. was unable to provide information for the questions addressed in our study, i.e. do the levels of E1A transcripts early in infection correlate with viral vDNA counts in the nucleus and is there variability in the transcription output from individual vDNAs within the same nucleus, or variability in how individual vDNAs within the same nucleus proceed into the replication phase. We hence do provide novel information, and do not consider this as a limitation of our paper.

We emphasize that population assays are done to attempt to understand molecular basis of a phenomenon by correlations. Instead, deep molecular insights require to-the-point-assays, in the case of transcription, single-molecule live cell assays at the level of single genes. Technically, we (and also the field) are not quite there yet.

Regardless, our study is a first step towards understanding transcription output of nuclear HAdV-genome at single-cell, single-genome levels. It has revealed insight that was not apparent from population assays. It is clear that the next step will be time-resolved live cell assays with simultaneous detection of transcription output, genome detection and transcription factor clustering on the genomic loci. With current technology the simultaneous detection of all these events is challenging, and requires the development of further technology.
Reviewer #2 (Evidence, reproducibility and clarity (Required)):

The authors show heterogeneity of AdV-C5 mRNA transcript quantity and dynamics in different cell types, which is regulated by the cell cycle phase and does not correlate to incoming viral DNA, using single molecule RNA FISH technologies and detection of incoming viral DNA by EdC labeling.

Major Comments:

The authors change the MOI used in their experiments (7 different MOIs are used throughout the paper) in a manner that appears randomly and without explanation. (54400 for Figure 1A, 1B, 3B, S3B; 37500 for Figure 1C; 23440 for Figure 2A, 2C, 5A; 13600 for Figure 1A, 1D; 36250 for Figure 3C, S3D; 11200 for Figure 4B; 23400 for Figure 6B). The authors should provide explanation, why these changes in MOIs are necessary.

AU: The MOIs given are theoretical MOIs, and essentially all figures indicate what was the actual MOI, that is, the real number of virus particles entering into the cells. This is beyond what is commonly provided in virology. It is essential, however, since MOI differs between different cell types. Therefore, we prefer to use the actual MOI as shown in Fig.1A, or we indicate the number of vDNAs that were delivered to the cells of interest.

Variable MOIs had to be used to ensure that different cell lines received comparable numbers of virions, in particular virus particle binding to and entering into the cells. Infection kinetics are different in different types of cells, but can be tuned by MOIs used. Furthermore, different virus preparations were used in the experiments and we performed analyses at different stages of the infection cycle. Due to all these different facettes provided by our experiments, it was impossible to choose one standard (theoretical) MOI for all the experiments.

The authors use mean fluorescence intensity of E1A probes per cell as estimate for viral transcript abundance for some of their experiments (Figure 1D, E, 3B), and count E1A punctae as measure for E1A transcripts in other experiments (Figure 2C, 3C, 5), without showing data, that these measures correlate. Problematic is hereby, that not all E1A punctae have the same signal intensity, as can be seen in Figure S1, which makes the estimation of the correlation of E1A punctae (= number of transcripts) and fluorescence intensity difficult. The authors should provide both (E1A punctae counts and estimation via fluorescence intensity) for at least one experiment, to prove, that the estimation of E1A transcript levels via fluorescence intensity is feasible.

AU: The quantification method had to be adjusted to the number of virus transcripts in the cell at the time of analysis. The best quantification method is segmentation and counting the individual fluorescent puncta per cell, but, as stated in the manuscript, this method does not accurately quantify the mRNA puncta from maximum projections of confocal or widefield image stacks when the number of puncta per cell exceeds ~ 200.

On the other hand, as shown in the quantification below, mean fluorescence intensity measurements per cell do not of course distinguish between cells having one vs. two mRNA puncta. Yet, as shown in the figure below, a relatively good correlation between puncta counting and fluorescence intensity measurements is achieved when cells have ≥ 10 transcripts per cell. Subsets of randomly picked images of the Fig.2C/Fig.5 dataset were included into the analysis ($r_s$ is Spearman’s correlation rank coefficient, approximate P<0.0001 for both; n is the number of cells analyzed). This figure is Fig. S1C in the revised manuscript.
p.15: "The nuclear E1A signals in AraC-treated cells were resistant to RNase A, but they were dampened by treatment with S1 nuclease (S6B Fig.)."
The authors make this statement based on (i) two completely different timepoints (12 h.p.i. for RNaseA treatment, 24.5 h.p.i. for S1 nuclease treatment) and (ii) in different clones of the A549 cells as stated in the methods section on p.21 (Two different clones of human lung epithelial carcinoma A549 cells were used in the study: our laboratory's old A549 clone (experiments shown in Fig. 1, Fig. 3B and S1 Fig., S3B and S3C Fig., S6A and S6B Fig., RNase A treatment) and A549 from American Type Culture Collection (ATCC, experiments shown in Fig. 2 and Fig. 5, Fig. 6, S2B Fig., S4 Fig., S5 Fig., and S6B Fig. S1 nuclease-treatment)). This makes it difficult to interpret, if the data is due to differences in the timepoints or cell types, or if it is due to binding of the E1A probe to single stranded vDNA.

AU: This is a fair criticism, thank you. We have replaced the RNase A figure S6B in the revised manuscript. A new RNase A experiment was repeated in ATCC A549 cells using the same infections conditions as with the S1 nuclease-treated cells.

Minor Comments:

p.4: "AdV are non-enveloped, double-stranded DNA viruses that cause mild respiratory infections in immuno-competent hosts, and establish persistent infections, which can develop into life-threatening infections if the host becomes immuno-compromised [reviewed in 6]."

Not all AdV cause respiratory diseases, the disease outcome of human AdV depends on the site of primary infection, which differs between the different AdV types.

AU: We have modified the text as follows: AdV are non-enveloped, double-stranded DNA viruses that cause mild respiratory, gastrointestinal or ocular infections...

p.7: The authors state, that "At the 17 h time point, about half of the cells had high numbers of protein VI transcripts, and most of them very high numbers of E1A transcripts.", however, the picture shown in Figure 1F shows a different phenotype, with low transcript levels of VI in E1A high cells and high transcript levels of VI in E1A low cells.

AU: This was perhaps a bit difficult to see in the overlay images since one has to distinguish between green and yellowish green. We have provided the individual channels along the overlay picture in Fig. S1D, and now it is clear that at 17h pi cells with high numbers of VI transcripts have also high numbers of E1A transcripts.

p.8: "This nuclear E1A signal is due to binding of the E1A probe to single-stranded vDNA in the replication centers (see below)."
The authors should state here, that due to the binding of the probes to the single stranded vDNA in the replication centers, the nucleus was excluded from the analysis for Figure 1F in late timepoints.

AU: We have modified the text according to the Reviewer’s suggestion. The text is now as follows: ‘Due to further studies (see below), we assume that this nuclear E1A signal represents binding of the E1A probe to single-stranded vDNA in the replication centers. Accordingly, the nuclear area was excluded when quantifying the viral transcripts per cell in late timepoints (Fig. 1F).’

Due to this time point the author cannot state that the E1A staining seen (Fig. 1F; indicated with white arrows) are replication centers; this is just an assumption, since there is no evidence in Fig 1 the author cannot be sure; the author should change the text: "taking the following experiments into account...", "due to further studies (see below)..... we assume that..."
AU: We have modified the text according to the Reviewer’s suggestion; see also the previous comment above.

p.8: The authors should mention the figure they refer to, since there is no E1B-55K staining in Fig. 1F

AU: The text has been modified as follows: Whereas other time points showed relatively few E1A, E1B-55K or VI puncta over the nuclear area (Fig. 1B, 1F, S1A Fig.), clustered nuclear E1A signals were apparent at 23 h.

p.9: Which test was used to calculate the additional p-values?

AU: As stated in the Material and Methods section or the figure legends, the p-values were calculated either by a permutation test using custom-programmed R-script (the code has been deposited on Mendeley Data along with other data associated with this manuscript), or by Kolmogorov-Smirnov test using GraphPad Prism. GraphPad Prism was also used to calculate Spearman’s correlation coefficients and the associated approximate p values. In the revised manuscript, we have added the following sentence into the Material and Methods section / Statistical analyses: Spearman’s correlation tests were done using GraphPad Prism.

p.10: For the experiment for the correlation of viral genomes per cell and E1A transcripts in HDF-TERT cells (Figure S2C), the MOI is missing in the description of the results, as well as in the corresponding figure legends.

AU: We have indicated the theoretical MOI (~ 4800 virus particles per cell) in the figure legend and in the Material and Methods section. The actual MOI, i.e. the actual number of virus particles entering into the cells, could not be determined due to the long (15 h) incubation time of virus inoculum with the cells, which in turn was required because these cells bind AdV-C5 rather inefficiently. However, between 1 and 32 EdC-labeled virus genomes were detected per cell nucleus at 22 h pi.

p.11: calculation of correlation? rs?

Why does the author combine S and G2/M phase? Fig. S3A show different values for the phases

AU: rs is the abbreviation for Spearman’s correlation coefficient, and, as indicated in the Material and Methods, we used GraphPad Prism to calculate the Spearman’s correlation coefficients.

Different methods to estimate cell cycle stages. DNA content method cannot separate S and G2/M with great confidence, whereas Kusabira Orange-hCdt1 and Azami-Green- hGeminin expressions in HeLa-Fucci cells allow more fine-tuned assessment of the cell cycle phases.

p.11: “Thus, the total intensity of nuclear DAPI signal can be used to accurately assign G1 vs S/G2/M stage to cells.”

The authors should also here refer to other papers, which showed that this correlation is feasible, as they did in the methods section (67. Roukos V, Pegoraro G, Voss TC, Misteli T. Cell cycle staging of individual cells by fluorescence microscopy. Nature protocols. 2015;10(2):334-48. Epub 2015/01/31. doi: 10.1038/nprot.2015.016. PubMed PMID: 25633629; PubMed Central PMCID:PMC6318798.), and maybe also refer to a newer paper which deals with this technique: Ferro, A., Mestre, T., Carneiro, P. et al. Blue intensity matters for cell cycle profiling in fluorescence DAPI-stained images.

AU: The integrated nuclear DAPI signal intensity is indeed a widely used method to assign cell-cycle stage to individual cells. We have added the second reference suggested by the Reviewer to the reference list for this method.

p.11: “Furthermore, when focusing on the highest E1A expressing cells, i.e. the cells with mean cytoplasmic E1A intensities larger than 1.5 × interquartile range from the 75th percentile, 71.9% of these cells were found to be in the G1 phase of cell cycle, whereas only 55.8% of cells in the total sampled cell population were G1 cells.”

The authors do not provide any reference to a figure within the manuscript or the supplements, which contains these data. Are these data not shown in the manuscript? AU: These values are calculated from the data shown in Fig.3B. The source data supporting findings of this study
(maximum projection images, excel files of the CellProfiler and Knime workflows) have now been deposited to Mendeley Data as stated in the Material and Methods / Data availability section of the revised manuscript and listed in Supplementary tables.

p.12; punctuation mistake; . instead of ,
To enrich G1 cells. AdV-C-5 (moi ~ 36250) was added. Why does the author switch between signal intensities and counting E1A puncta per cell (limited to 200) in the different experiments to illustrate accumulation of E1A transcripts?

AU: The same answer as above: the quantification method had to be adjusted to the number of virus transcripts in the cell at the time of analysis. The best quantification method is segmentation and counting the individual fluorescent puncta per cell, but, as stated in the manuscript, this method does not accurately quantify the mRNA puncta from maximum projections of confocal or widefield image stacks when the number of puncta per cell exceeds ~ 200. On the other hand, as shown in the quantification in the new S1C Fig., mean fluorescence intensity measurements per cell do not of course distinguish between cells having one vs. two mRNA puncta, but a relatively good correlation between puncta counting and fluorescence intensity measurements is achieved when cells have ≥ 10 transcripts per cell.

p.14: "For E1A (or E1B-55K), we did not detect transcriptional bursts with bDNA-FISH probes on nuclear vDNAs, either prior to or after accumulation of viral transcripts in the cell cytoplasm.” The authors do not provide any reference to a figure within the manuscript or the supplements, which contains these data. Are these data not shown in the manuscript? AU: This statement is based on hundreds of images we have analyzed during the course of the study. It is impossible to show all of these images, so in principle, this is “data not shown”. We have modified the text as follows: With hundreds of images analyzed, we never unambiguously detected transcriptional bursts with E1A (or E1B-55K) bDNA-FISH probes on nuclear vDNAs, either prior to or after accumulation of viral transcripts in the cell cytoplasm.

p.14: space between number and %
AU: Thank you for pointing this out. It has been corrected.

p.15: "This is was also seen in AdV-C5-EdC-infected cells” should be changed to "This was also seen in AdV-C5-EdC-infected cells”
AU: Thank you for pointing this out. It has been corrected.

Fig. 1B:
- figure legend does not indicate how cells were stained
- also no description in the continuous text
- which E1A transcripts are stained? all? 12S? 13S?
AU: The first sentence in Results section states that “We used fluorescent in situ hybridization (FISH) with probes targeting E1A, E1B-55K and protein VI transcripts followed by branched DNA (bDNA) signal amplification to visualize the appearance and abundance of viral transcripts in AdV-C5-infected A549 lung carcinoma cells.” Furthermore, the legend to Figure 1 starts with the title “Visualization of AdV-C5 E1A, E1B-55K and protein VI transcripts in infected cells by bDNA-FISH technique”, and the legend to Fig.1B mentions that “cells were stained with probes against E1A and E1B-55K mRNAs or E1A and protein VI mRNAs”. We are of the opinion that this is enough information to understand the figures.

The main text to Fig.1 also states that “The E1A probes covered the entire E1A primary transcript region and thus all E1A splice variants. The temporal control of E1A primary transcript splicing and E1A mRNA stability give rise predominantly to 13S and 12S E1A mRNAs at 5 h pi (references)”.

Fig. 1D:
- difference in accumulation of viral transcripts is not that visible as in IF staining (Fig. 1B; Fig. 1S); Fig. 1 or S1 Fig. do not show IF staining but signals from FISH.
- graph does not show any difference between E1A and E1B-55K
AU: The y-axes values in Fig. 1D graph are arbitrary units and thus E1A and E1B-55K graphs are not directly comparable to each other. We have included into the revised manuscript S1B Fig., which shows quantification of E1A and E1B-55K fluorescent puncta per cell at the 5 h pi; the difference between E1A and E1B-55K was statistically significant.

Fig. 1F:
-figure legend does not fit with labelling of IF images and continuous text
-description says 22 h, while IF labeling and text (p. 7, last lane) mentions 23 h
AU: The figure annotations state the time of analyses as total time after virus addition to cells, whereas text stated the time of analyses as x h post virus removal since we wanted to stress that the input virus was incubated only for 1 h with the cells. However, Reviewers found this confusing, so we have changed the text in the revised manuscript so that time of analysis is stated as total time after virus addition to cells (as in the figure annotations). Only in the Material and Methods section we maintain the original 1 h + x h statement for the time of analysis.

Fig. 2A:
–figure legend: lane 5
Punctuation wrong: azide-Alexa Fluor488. Alexa Fluor647
AU: Thank you for pointing this out. It has been corrected.

Fig. 4A:
–difficulties to understand
–author stated that promoter-driven EGFP expression is clearly dominated by G1 cells for E1A and by S/G2/M cells for CMV, however this is not clearly visible in the graph
–no severe differences visible between CMV-eGFP and E1A-eGFP
–author should include numbers for quantification and statistical calculations to illustrate the differences
AU: In the highest GFP expression bin, CMV-eGFP expressing cells have 43% cells in G1 and 50% in S/G2/M (n=2149). In comparison, E1A-GFP expressing cells have 58% cells in G1 and 35% in S/G2/M (n=2258). The difference in G1 cells in the highest eGFP bin is statistically significant (p<0.00001, Two Proportion Z-test). This quantification has been added to the main text, as well as to the figure.

Fig. 4B:
–amount of E1A protein levels calculated via IF (signal intensities)
–immunofluorescence is not a suitable tool for protein quantification
AU: It is true that not all antibodies are suitable for IF (or for Western blot), and we cannot be certain that the monoclonal anti-E1A antibody used by us detects all E1A forms with different post-translational modifications with equal efficiency. However, IF is a widely accepted method to estimate protein levels in the cell, especially if the proteins like E1A accumulate in the nucleus (makes segmentation of the signal easy) and give a rather uniform nuclear staining pattern.

Fig. 5:
–in A. it is stated, that E1A bDNA-FISH is not suitable, since it is too short to be detectable. However, in B E1A bDNA-FISH is used. is there a difference?
–according to the method part just one E1A mRNA was used for the assays, why is it then not possible to use that one in Fig. 5A?
–explanation of the procedure and the experiment is very confusing
AU: The Reviewer probably refers to Fig. 6 here, not to Fig. 5. The E1A introns are short (about 100 bases) and cannot be picked up with bDNA FISH probes. In Fig. 6B we were using the E1A bDNA-FISH probes, which were made against the AdV-C5 genome map positions 551-1630 to detect vDNA single strands of the E1A region and these single strands were long enough to be picked out by our E1A probes.

Fig. S6B:
–authors want to show that it is RNase-insensitive, but S1 nuclease-sensitive
–two different A549 cell clones and two different time points are used for the treatments
AU: This is a fair criticism. We have replaced the RNase A figure in S6B Fig. in the revised manuscript. The new RNase A experiment was carried out in ATCC A549 cells using the same infections conditions as with the S1 nuclease-treated cells.

Material and Methods:
- headings do not indicate which methods are explained
- no clear structure

AU: We have made minor changes to the headings of Material and Methods section. We have first explained in detail the bDNA-FISH method, but otherwise the order is according to the order of the figures.

Reviewer #2 (Significance (Required)):

highly significant manuscript
very important for the virology field

my research topics are human adenoviruses and their replication cycle

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

Summary:
Soumalainen et al have studied adenovirus viral gene expression and replication at a single-cell level. They explore the extent of correlation between incoming genome copy number and early gene expression and progression into the late phase, revealing substantial variation between cells in the numbers of E1A transcripts (the first gene expressed upon infection) that is not explained by differences in the numbers of viral genome templates in the cells. They also explore the relevance of cell cycle stage to this variability and show a positive correlation between G1 cell cycle stage and higher levels of gene activity, which explains at least part of the variation.

To form these conclusions they have applied new methods to visualise and quantify single molecules of nucleic acid in single cells. The experiments are all carefully and fully described with full detail of materials. Overall the manuscript is well written and easy to follow.

Major comments:

All of the experiments appear to be done with rigour and their results reported with due regard to statistical significance etc.

My major concern though is that they have been done, perhaps out of necessity to get detectable signals, at very high multiplicities of infection. A well-accepted standard to achieve infection of all cells in a culture is an MOI of 10 infectious units per cell. Even this is acknowledged not to represent the biology of natural infection and it is striking that, where technically feasible, lower MOI studies are more revealing of how a virus actually works. Here, the authors have used counts of particles rather than infectious units to determine MOI and for Ad5, the particle/pfu ratio is typically 20-100. Their MOIs though are 13,000 - 50,000 per cell, implying an infectious MOI of at least 130 for their A549 experiments, which are known to be readily infected by Ad5 from other work.

AU: Unlike common experiments done by others, we used a synchronized infection and removed the input virus after 1h incubation at 37°C. This type of infection initiation requires high input virus amounts, as opposed to studies in which the virus inoculum is incubated with cells for several hours/days, as is typically done in studies determining the infectious or plaque forming units in virus inoculum. Hence, the MOI used by others involved incubation of inoculum with cells over extended periods of time, and they cannot be compared to our pulsed infection conditions.

Although the calculated theoretical MOIs (physical particles/cell) were high in our experiments, only 0.1% - 0.2% of input virus particles bound to cells during the 1h incubation period (Fig. 1 A;
this estimation is based on the ratios between Median values for the number of cell-associated viruses vs input virus numbers).

Furthermore, in the experiment described in Fig.2C (correlation of E1A transcripts per cell vs. viral genomes per cell), 42% of analyzed cells had ≤ 5 viral genomes/cell and 27.5% of analyzed cells had between 6-10 viral genomes per cell. Please note, that these are not high numbers.

The input virus amounts used were selected this way, because we aimed at getting a broader view of how virus transcription at early phases of infection responds to a varying number of virus genomes delivered to the nucleus. Therefore, we did not limit the analyses to a situation with 1 or less than 1 virus particles/genomes per cell.

In addition, the analyses of how cell cycle phase impacts the initiation of virus gene expression requires a relatively short time between virus inoculation and time point of analysis (i.e. a rather high MOI). Otherwise, as also pointed out by the Reviewer, the cells could have experienced more than one cell cycle phase during the duration of the experiment. Furthermore, although the initial natural infection probably starts with a very low MOI, the second round of infection is a high MOI infection due to a large number of progeny virus particles released from an infected cell.

Surprisingly, the authors do not see intracellular vDNA copy numbers that are fully reflective of this high MOI, with median intracellular vDNA of 75/cell at the highest MOI. The authors should consider how the population distribution of vDNA/cell does or does not fit the predicted Poisson distribution. Nonetheless, at these high copy numbers/cell, there must surely be a risk that the variation in gene expression activity arises stochastically, out of competition between genomes for essential transcription factors. Given that multiple cellular factors are each required for E1A transcription, high genome copy numbers could actually inhibit E1A expression relative to cells with more modest copy numbers because limited supplies of individual factors are recruited to different viral genome copies.

AU: The “discrepancy” between theoretical MOI and the actual observed number of cell-associated virus particles or cell-associated virus genomes is explained above. Furthermore, we would like to point out that we have directly estimated the number of virus particles bound to cells with the input virus amounts used, something that is usually not done in other studies.

It is indeed theoretically possible that high nuclear genome numbers could lead to inhibition of transcription due to competition for limiting essential host factors. However, if we included only cells with ≤4 vDNA molecules per nucleus into the analysis (total number of cells analyzed was 258), then Spearman’s correlation coefficient for vDNA per nucleus vs E1A mRNAs per cell was 0.186 (p=0.0027). Thus, this would not support the notion that cells with moderate nuclear vDNA copy numbers would have a better correlation between the nuclear vDNA copies vs E1A mRNA counts per cell.

The vDNA/cell in Fig.2C does not fit predicted Poisson distribution, var/mean=9.129.
It is important for the analysis of correlation of gene expression with cell cycle that the virus has not, at the time point analysed, already perturbed the cell cycle (a well-known effect of infection) which the authors document in Suppl Fig3B. To my eye, the G1 peak in infected cells is somewhat narrower than in the control while the S/G2 bump is a little greater. The % of cells in each of the two gates needs to be shown to support the conclusion.

AU: In non-infected sample G1 = 54.63% and S/G2/M = 45.37%, in infected cells G1 = 51.4% and S/G2/M = 48.6%. We have added this information into the S3B Fig.

Turning to the experiments documenting a correlation between E1A expression and cell cycle stage, the authors interpret their findings in terms of the stage the cells are at when the analysis was done (G1 stage cells have more E1A transcripts). The key experiment (Fig 3B) is analysed at only 4 h pi, so substantial progression from G2/M back to G1 after virus addition can probably be discounted, but the point should be discussed. The authors also use release from G1 in another cell line to support their argument that G1 supports higher levels of E1A expression (Fig 3C). Here, they elect to exclude all cells with fewer than 50 E1A transcripts from their analysis.

The reason for this is completely obscure and isn't obviously justified; conceivably it could bias the outcome of the experiment. At minimum, this decision needs to be carefully explained; ideally, the full data set should be used.

AU: Fig.3B: As suggested by the Reviewer, we have added to the main text the following explanation: “We used a high MOI infection (median 75 cell-associated virus particles, Fig. 1A) in order to achieve a rapid onset of E1A expression so that the time between virus addition and analysis was short. Thus, it is not expected that a substantial number of cells would have changed their cell cycle status during the experiment.”

Fig.3C: We show the results also from the full data set of infected cells, i.e., cells with ≥ 1 E1A puncta in S3D Fig. We excluded the cells without zero E1A puncta because with these cells it is impossible to know whether they received no virus or whether E1A transcription had not yet started. Permutation test indicated that the difference between the starved+starved and starved+FCS is statistically significant even in this case. Because both samples are dominated by cells with low E1A counts, we log-transformed the E1A values for the box plot figure.

The authors note the highest level of E1A activity (as opposed to RNA) was in G1/S cells and suggest that high E1A cells advance preferentially into S. Whilst in line with the literature that E1A promotes progression into S, an alternative explanation is simply that there is a time lag between RNA accumulation and protein accumulation, during which progression through the cycle would be expected.

AU: This is a valid point, and we have modified the text as follows: “… which could reflect the advancement of high E1A expressing cells into S-phase. However, considering the time between
virus addition and analysis (10.5 h), we cannot exclude the possibility that the observed G1/S preference is at least partly due to time-dependent progression of G1 cells to G1/S.”

Minor comments:
Fig 1 and elsewhere. Given that the 1 h incubations with virus were done at 37 C, the convention would be to include this period in the time post-infection at which harvest / fix time points are quoted. There is inconsistency between text and legend with 12 h pi being sometimes represented as 11 h after virus removal; this is an unnecessary confusion.
AU: We have modified the text so that hours pi always include the 1h incubation with the input virus. Only in the Material and Methods section we kept the original 1h virus binding – fixing at xh post virus removal.

Results description prior to the ref to Fig 1B: unclear what this is supposed to mean. AU: We have now slightly modified the first paragraph of the Results section. We mention the benefits of the bDNA signal amplification method and explain the experimental set up, i.e. that the input virus was incubated with the cells only for 1h. We also justify why we used a short incubation for the virus inoculum.

Fig 4A: provide % of cells in each gate in each histogram.
AU: In the highest GFP expression bin, CMV-eGFP expressing cells have 43% of cells in G1 and 50% in S/G2/M. In comparison, E1A-GFP expressing cells have 58% of cells in G1 and 35% in S/G2/M. This has been added to the figure, and it is also mentioned in the main text. Furthermore, we added to the text the results from Two Proportion Z-test to show that the proportion difference of G1 cells in the highest bin was statistically significant (p<0.000001).

Fig 5: bottom right panel x axis label is wrong
AU: Thank you for pointing out this. This has been corrected.

In the presentation of Fig 6, it would be much clearer for the reader if the detected replication foci (ssDNA detected as E1A puncta) were referred to as something other than E1A puncta. There is too much scope for confusion with the earlier experiments in which E1A RNA was detected.
AU: We agree. In the revised manuscript, we refer to these puncta in the text as E1A ssDNA-foci.

Reviewer #3 (Significance (Required)):

The study represents the application of state of the art single-molecule visualization techniques to an as yet not understood aspect of virus infection. That said, there is prior experimentation in this area, which the authors fully acknowledge and build upon. The new work is largely descriptive, in that it reveals very clearly the discrepancy between genome copy number and amounts of mRNA without seeking to explain these, beyond the cell cycle analysis. Whilst there is a better correlation between vDNA number and transcript once the data are stratified by cell cycle stage, it is still not strong (Fig 5), indicating that other substantial contributing factors remain to be described.

The work will be of interest certainly to adenovirologists, but also to others who study virus infections - particularly nuclear-replicating DNA viruses such as herpesviruses - where similar considerations are likely to apply.

Expertise: adenovirus; gene expression; virus-host interactions; molecular biology

First decision letter

MS ID#: JOCES/2020/252544

MS TITLE: Cell-to-cell and genome-to-genome variability of Adenovirus transcription tuned by the cell cycle
AUTHORS: Maarit Suomalainen, Vibhu Prasad, Abhilash Kannan, and Urs F Greber
ARTICLE TYPE: Research Article

Thank you for sending your manuscript to Journal of Cell Science through Review Commons.

I am happy to tell you that your manuscript has been accepted for publication in Journal of Cell Science, pending standard ethics checks and minor text additions suggested by one reviewer. As we discussed by e-mail, I think it would be helpful to include more details such as scripts for the pipelines.

Reviewer 1

Advance summary and potential significance to field

Please note that I reviewed a preliminary version of this manuscript when submitted to Review Commons (Reviewer #3). Suomalainen et al have studied adenovirus viral gene expression and replication at a single-cell level, seeking to understand the diversity in outcome between individual cells and viral genomes. They explore the extent of correlation between incoming genome copy number and early gene expression and progression into the late phase, revealing substantial variation between cells in the numbers of E1A transcripts (the first gene expressed upon infection) that is not explained by differences in the numbers of viral genome templates in the cells as they show no correlation between E1A txt #s and intracellular viral genome #s. They also explore the relevance of cell cycle stage to this variability and show a positive correlation between G1 cell cycle stage and higher levels of gene activity, which explains at least part of the variation. Lastly they show that individual viral genomes progress to replication with considerably different rates within the same nucleus.

To form these conclusions they have applied state-of-the-art methods to visualise and quantify single molecules of nucleic acid in single cells. Although there is a pre-existing literature in this area, the authors’ approach has produced significant new insights, as well as indicating how these questions might be addressed in other systems, and thus does display significant novelty.

The experiments are all carefully and fully described with full detail of materials. Overall the manuscript is well written and easy to follow, and the text has been further improved in response to preliminary review.

Comments for the author

The principal concern I (and other reviewers) had on preliminary review was over the very high multiplicities of infection used for all experiments, along with variation in that multiplicity between experiments. The authors have responded by explaining clearly their approach within the paper, and have justified it as being necessary to establish rapid synchronous infection over a short incubation (1 hr). They have carefully measured the actual infecting multiplicity in their experiments (genomes/hexon puncta in newly infected cells) and I accept this is a firmer foundation for conclusions about (lack of) correlation between single-cell MOI and progress of infection than is provided in any standard study. What is less reasonable is their proposition that standard experiments with more modest MOIs require much longer incubations with virus to achieve infection which is therefore asynchronous among the cells of the population so as to confound the kind of analysis they wished to do: at least in tumour cell lines such as A549, this is not the case in my experience, though normal fibroblasts are indeed much less easy to infect.

In response to a query concerning the possible competition between large numbers of genomes for transcription factors as a potential explanation for the lack of correlation between E1A transcripts and vDNA, they've indicated in response that restricting the analysis only to cells with between 1 and 4 vDNA copies did not improve the correlation. I would like to see this comment added to the m/s on p9 as it goes a long way to answering the concern that the high multiplicity is a factor in the results.

All the other points have been effectively addressed and I have no further points to add.