THE
EMBO
JOURNAL

Live-cell imaging reveals the interplay between transcription factors, nucleosomes, and bursting

Benjamin T. Donovan, Anh Huynh, David A. Ball, Heta P. Patel, Michael G. Poirier, Daniel R. Larson, Matthew L. Ferguson and Tineke L. Lenstra

Review timeline:

Submission date:	28th Sep 2018
Editorial Decision:	6th Nov 2018
Revision received:	15th Mar 2019
Editorial Decision:	5th Apr 2019
Revision received:	12th Apr 2019
Accepted:	25th Apr 2019

Editor: Anne Nielsen / Hartmut Vodermaier

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. The original formatting of letters and referee reports may not be reflected in this compilation.)

1st Editorial Decision 6th Nov 2018

Thank you for submitting your manuscript for consideration by the EMBO Journal. It has now been seen by three referees whose comments are shown below.

As you will see from the reports, our referees all highlight the importance and quality of your data but also raise a couple of points that will need to be addressed before they can support publication of the manuscript here.

For the revised manuscript I would particularly ask you to focus your efforts on the following concerns:
-> Please follow the suggestion from ref #1 to look at the Gal80 dynamics during Gal4-induced transcription. I also think point #2 from this referee would be a great addition to your study but I realise that it may be technically challenging to do and I'd be happy to discuss the details and what you could include to address this point.
-> Refs #1 and #3 both ask for a more extensive discussion of the model to acknowledge the complexity in transcriptional control.
-> Please include a more stringent size fractionation when analyzing the MNase data as requested by ref #2.

Given the referees' overall positive recommendations, I would like to invite you to submit a revised version of the manuscript, addressing the comments of all three reviewers. I should add that it is EMBO Journal policy to allow only a single round of revision, and acceptance of your manuscript will therefore depend on the completeness of your responses in this revised version.

REFEREE REPORTS:

Referee #1:

Report on manuscript "single-molecule imaging reveals the interplay between transcription factors, nucleosomes, and transcriptional bursting", by Benjamin Donovan et al.

In this manuscript, the authors use advanced imaging techniques to explore the mechanism of bursting of the galactose-inducible genes in yeast. By combining single molecule microscopy of RNA and of the transcription factor Gal4, the authors reach several conclusions: (i) that the dynamic of Gal4 binding onto the UAS sequences controls bursting, and in particular that Gal4 dwell time sets burst duration; (ii) that galactose induction increases burst frequency but not burst size.

Overall, the study is technically sound, and by simultaneously tracking single binding events of Gal4 on a reported gene, together with gene activity using PP7, the authors realize a technical tour-de-force that allow them to directly show that a Gal4 binding event on chromatin produces a burst of promoter activity. A few studies have attempted to do this, but so far relied on indirect correlations. The authors thus bring an important result and I fully support publication in EMBO J., provided the comments below are addressed.

1-The dynamic of Gal4 binding on DNA vs the dynamic of Gal80 on Gal4.
The authors try to prove that the bursting of the Gal genes is controlled by the binding/unbinding of Gal4 on DNA. This is somehow counterintuitive because decades of biochemical studies have shown that galactose does not regulate the binding of Gal4 on DNA (it is always bound), but controls the binding of the repressor Gal80 on Gal4. I understand that "always bound" in biochemical terms can still mean a very dynamic system. Nevertheless, the big unknown that is lacking to fully prove the model is the dynamics of Gal80 on Gal4 (if possible in different sugar condition). Indeed, if one imagines that the binding of Gal80 on Gal4 is much more stable than the binding of Gal4 on DNA, then when Gal4 binds DNA, it is either an active molecule or an inactive molecule (ie bound by Gal80), and the situation is simple: an active molecule produces a burst, while an inactive molecule produces nothing, and the interpretation of the authors makes perfect sense.
However, if the dynamic of Gal80 on Gal4 is similar or faster than that of Gal4 to DNA, then one expect that a burst might be created by either the binding of a free molecule of Gal4, or by the release of Gal80 from Gal4. Likewise, the burst duration may not only be controlled by the rate of release of Gal4 from DNA, but also by the rate of binding by Gal80 to Gal4. To prove their point, it is thus essential that the authors investigate the in vivo dynamics of Gal4 binding to Gal80. This could be done in several ways. For instance, the concentration of Gal80 could be varied, or the life-time of the complex could be measured directly. For this, Gal4 could be for instance fused to Laci and forced to bind a LacO array, while the spt technique of the authors could then be used to measure the dwell time of Gal80 on the array.

2-Binding of Gal4 to the UASmut.
The authors make a lot of nice experiments in Figure 1 and 2 to decrease the dwell time of Gal4 to DNA and measure the consequences on gene activity. It is therefore a pity that the live-cell measurement of Gal4 binding is made on the Gal10 gene and not on Gal3 (Figure 3). To prove their point that the dwell time of Gal4 indeed controls burst duration, I encourage the authors to perform the live cell measurements of Gal4 binding to chromatin on Gal3 and to compare the WT and mutant UAS. This would make their point much stronger. The authors allude that this is technically challenging. In this respect, a system that might useful is the one described in PMID: 19651897, where a small gene array is used next to a LacO array, and allow to significantly amplify the signal of Gal4 binding.

3-Figure 6.
I am not convinced by the measurements of burst duration presented here. Indeed, as seen in Figure 6A and as the authors state, the "bursts" (in the sense of the ON periods of the promoter) overlap, and the measurement of the burst size become somehow arbitrary. I am also not convinced by the autocorrelation analysis. To my knowledge, there is currently no published methods for how to fit an MS2/PP7 autocorrelation curve in the case of a two-state promoter model, and the assignment of the "burst" duration thus also appears not to be on very solid grounds (see for instance fig 5D where the two linear fits are not very good). Furthermore, the dwell-time of the nascent RNA appears to be in the same order as the Gal4 dwell time, and as the delay between bursts, complicating the situation. I

am also worried that other processes might be at play that make the analysis of the correlation even less reliable. Are the autocorrelation curves invariant with respect to the duration of the acquisition ?

Finally, when the gene is fully induced, all the four binding sites of the Gal10 promoter are occupied, presumably by active Gal4 molecules, and is well possibly that bursting then becomes controlled by other factors, such as the lifetime of the PIC (which may be independent of Gal4 dynamics). One way to determine this would be to repeat the experiment with the UASmut and see whether it changes anything, both in terms of mature mRNA produced and dynamics. This figure may also be simplified or dropped if the authors wish to, since at present it does not convey essential conclusions.

Minor comments:

-Figure 1: as a control, it is important to quantify the mRNA levels produce by these promoters to measure their relative strength.

-Figure 3: as a control, the authors should show that the Gal4-Halo is fully functional and does not lead to alteration of promoter regulation.

-Figure 4F: can the measurements be done both in galactose and raffinose ?

-The authors could indicate on each panel the sugar conditions used, to facilitate reading.

-In Figure 7, the authors may want to include Gal80 in their model.

-For the orbital tracking, I was not sure of what happens when the transcription site goes off. Does the system keeps the same Z position waiting ?

Referee #2:

This paper uses some advanced microscopy to visualize events at a single gene locus in vivo. Monitoring both Gal4 activator and the transcript, the authors reach several conclusions. Not too surprisingly, Gal4 affinity correlates with dwell time in vivo and in vitro. The presence of nucleosomes decreases Gal4 dwell time (see comment below). Finally, looking at Gal4 binding versus the transcription burst duration and frequency, it appears that transcription is limited to the window when activator bound. Therefore, weaker Gal4 binding not only reduces Gal4 dwell time but also burst duration. In contrast, increased gene expression by galactose induction increases burst frequency, but not duration. The work is convincing and the observations provide interesting new insights into how we should think about transcription activation mechanisms.

I have some specific comments, which are mostly things that need some clarification:

p5. It would be helpful to know if "one GAL4 UAS" in the Gal3 promoter means it contains a single Gal4 binding site, or it's one UAS made up of multiple Gal4 binding sites.

p5 The authors apply a threshold to the PP7 signals in Figure 1, the details of which are described in the methods section. As this seems to be based on the standard deviation in the background, is there some way to be sure this doesn't filter out single transcription events?

p6. If the UASmut only makes one transcript at a time, can that still be called a "burst"? This is a question of semantics, but producing single transcripts is also used as a distinguishing property of "constitutive" genes, so that definition becomes less useful if there can be single transcript "bursts".

p6. I don't think you want to say that "competitor DNA was more effective in removing Gal4 from the labeled DNA". This is an equilibrium assay, so the competitor simply binds Gal4 that dissociates from the labeled DNA. The way it's written makes it sound like there's an active removal process.

Fig 2D - I recommend adding the red fluorophore into the schematics at the bottom (similar to 2A). Based on the picture, this must be in the middle of the probe?

MNase expt (Fig 3). Defining 95-225 nt fragments as nucleosomal size seems overly broad, especially since there's still debate about calling this type of weaker MNase protections as "fragile nucleosomes". Are the results any different if you set the bottom limit should be about 140-150? Other protein complexes can give footprints below that length. How can we be sure there's not something else there, perhaps even Gal4 plus other activators and coactivators?

p10 There's a sentence about the 152s duration of Gal10 transcription being compatible with a previously reported dwell time. It's unclear: dwell time of what? If this refers to the transcript appearance, that's a little confusing because this paper usually uses "on time" or "duration" to describe that time period, while "dwell time" is usually used for Gal4 binding. The authors may want to check throughout the paper to make sure they are using terms consistently so readers don't get confused.

p10 I don't follow the logic of this conclusion:
"overlap between Gal4 binding and GAL10 RNA appearance agrees with our model that several RNA polymerases initiate while Gal4 is bound to the promoter"
Why couldn't there also be overlap for single transcription events? I think this point needs some further explanation to make the argument clear.

Referee #3:

For a variety of technical reasons, linking the kinetics of transcription factor binding to the transcriptional output from a gene locus through real-time observation using optical microscopy has proved difficult. However, such knowledge is of fundamental importance in dissecting the mechanisms linking the two process. Multiple recent reports of transcriptional system with transient bursting behaviors that are not at steady-state highlight the importance of overcoming this challenge. Benjamin et al. have presented a comprehensive study on how the binding kinetics of Gal4 impact the subsequent transcriptional output of genes under its regulation. They combined in vitro and in vivo imaging approaches, supplemented with established genomics techniques where appropriate to establish a broader context in vivo, to characterize how different aspects of the GAL3 and GAL10 systems impact the kinetics of Gal4 binding, and, subsequently, how they change the behavior of transcription. Their observations lead to important conclusions, especially that the kinetics of Gal4 binding in their system is sufficient to produce bursting behavior without necessitating additional higher-order organizations in the nucleus. Their discovery of an upper-limit of Gal4 binding affinity for improving transcriptional output in vivo that is still significantly weaker than the consensus binding sequence for Gal4 is likewise intriguing, indicating that there are more parameters that constrain the evolution of the system beyond binding kinetics.

Overall, the experimental design is well thought out and the execution is also of good quality. Control experiments are done to limit artefacts when correlating signals and the limitations of each technique are also clearly explained and accounted for. While the experiments are conducted in yeast cells, the results and conclusions are sufficiently fundamental that they improve our basic understanding of transcriptional bursting and should remain applicable in more complex systems. If the authors could address the following questions/issues in the manuscript, I would recommend it for publication.

1. The authors stated the they chose the GAL10 system for the experiments in the second half of the manuscript but used the GAL3 system for the first half. Is there a reason they specifically chose GAL3 for the experiments presented Figures 1 and 2? Does GAL10 contain other elements that could complicate kinetic analysis? If not, could they compare the two systems when their kinetics are measured in the manner of Figure 1 panels E, F, and I? Additionally of interest: Do the authors have information on where GAL10 wt would fall in Figure 2K compared to GAL3 wt?

2. On page 8, line 19: the authors stated that they "observed single Gal4 molecules in the nucleus." Could the authors explicitly state their evidence for such a claim? Does the intensity over time

change in discrete steps?

3. Page 8, lines 30-32 and Figure EV4 B and C: The authors stated that H3 does not exhibit short-binding population. In Figure EV4 panels B and C, the authors showed a single-exponential fit for the H3 population. However, there seems to be a tail at longer time (>30~50 s) in the distribution that is not accounted for by the fit. If the H3 population is fit using a bi-exponential distribution, is there a significant population that is even longer lived? Is the "fast component" of H3 more stable than the "slow component" of Gal4? The captions in panel D suggests that the author did such a fit with H3.

4. Page 8 the bottom half of the last paragraph, Figure 4 and EV4: What are the relative populations of the fast and slow components? Do the relative populations change between galactose and raffinose or does only the slow dwell time change?

5. Page 9, the paragraph spanning lines 12-23: The authors stated that H3 binds on much longer time scales and they used it as a control of the maximum measurable dwell time. However, for the 200 ms tracking experiments, H3 in Figure EV4D has a slow component that lasts for a shorter time than Gal4 in the presence of galactose. What is the reason for this discrepancy?

6. Page 11 lines 12-14: Have the authors observed sufficient differences in photobleaching rates between orbital-tracking and time-lapse experiments that led to their suggestion that it could be a reason for the difference in the measured burst durations?

7. In general about the discussion and Figure 7: The authors provided a very simplified figure as a summary of their findings and for possible mechanisms that they hypothesized based on their results. The authors should expand the figure such that it provides a more comprehensive summary of the salient properties of Gal4 driven regulation as explored in this manuscript and their hypothesized mechanism of action. For example, multiple RNA PolII recruited when TF is bound, termination of burst upod TF dissociation, how TF binding and burst kinetics are affected by various factors (binding site affinity, nucleosomes, galactose), etc.

Minor issues:
1. Page 8, lines 21-23: The authors should cite the methods section describing how H3 movement is used to select a threshold between frames used to determine if a particular Gal4 track is bound. It was unclear how this was done and what the threshold is when I first read this section and it was unclear where I could find the answer.

2. Page 9 line 19: "Figures S4B-D" should be "Figure EV4B-D"

1st Revision - authors' response 15th Mar 2019

Response to reviewers

Referee #1:

Report on manuscript "single-molecule imaging reveals the interplay between transcription factors, nucleosomes, and transcriptional bursting", by Benjamin Donovan et al.

In this manuscript, the authors use advanced imaging techniques to explore the mechanism of bursting of the galactose-inducible genes in yeast. By combining single molecule microscopy of RNA and of the transcription factor Gal4, the authors reach several conclusions: (i) that the dynamic of Gal4 binding onto the UAS sequences controls bursting, and in particular that Gal4 dwell time sets burst duration; (ii) that galactose induction increases burst frequency but not burst size.

Overall, the study is technically sound, and by simultaneously tracking single binding events of Gal4 on a reported gene, together with gene activity using PP7, the authors realize a technical tour-de-force that allow them to directly show that a Gal4 binding event on chromatin produces a burst of promoter activity. A few studies have attempted to do this, but so far relied on indirect correlations.

The authors thus bring an important result and I fully support publication in EMBO J., provided the comments below are addressed.

1-The dynamic of Gal4 binding on DNA vs the dynamic of Gal80 on Gal4.
The authors try to prove that the bursting of the Gal genes is controlled by the binding/unbinding of Gal4 on DNA. This is somehow counterintuitive because decades of biochemical studies have shown that galactose does not regulate the binding of Gal4 on DNA (it is always bound), but controls the binding of the repressor Gal80 on Gal4. I understand that "always bound" in biochemical terms can still mean a very dynamic system. Nevertheless, the big unknown that is lacking to fully prove the model is the dynamics of Gal80 on Gal4 (if possible in different sugar condition). Indeed, if one imagines that the binding of Gal80 on Gal4 is much more stable than the binding of Gal4 on DNA, then when Gal4 binds DNA, it is either an active molecule or an inactive molecule (ie bound by Gal80), and the situation is simple: an active molecule produces a burst, while an inactive molecule produces nothing, and the interpretation of the authors makes perfect sense.
However, if the dynamic of Gal80 on Gal4 is similar or faster than that of Gal4 to DNA, then one expect that a burst might be created by either the binding of a free molecule of Gal4, or by the release of Gal80 from Gal4. Likewise, the burst duration may not only be controlled by the rate of release of Gal4 from DNA, but also by the rate of binding by Gal80 to Gal4. To prove their point, it is thus essential that the authors investigate the in vivo dynamics of Gal4 binding to Gal80. This could be done in several ways. For instance, the concentration of Gal80 could be varied, or the life-time of the complex could be measured directly. For this, Gal4 could be for instance fused to Laci and forced to bind a LacO array, while the spt technique of the authors could then be used to measure the dwell time of Gal80 on the array.

As reviewer 1 points out, Gal4 is repressed by Gal80 and that in principle the binding kinetics of Gal80 to Gal4 could control the activity of Gal4 and thereby the burst duration. However, the experiments linking Gal4 dynamics to burst duration have all been performed at full induction conditions (2% galactose), and Gal80 is expected to be fully dissociated from Gal4 under these conditions (Jiang *et al*, 2009)*. To verify that Gal80 does not influence transcriptional bursting in these conditions, we followed the reviewer's advice and deleted GAL80. smFISH experiment show that the nascent transcript distributions of GAL10 and GAL3 are similar in the presence or absence of Gal80 (Figure EV4). In full induction conditions, it is therefore highly unlikely that Gal80-Gal4 binding contributes to the control of burst duration. In addition, we find the burst duration does not change across different galactose concentrations (with different levels of Gal80 activity/concentration), suggesting that Gal80-Gal4 kinetics does not contribute to the burst duration. The reduction of the burst frequency in lower galactose concentrations suggests that bound Gal4 is either active (not bound by Gal80 and producing a burst), or inactive (repressed by Gal80). We have added these smFISH results (Figure EV4) and a section describing the role of the Gal80-Gal4 interaction to the manuscript (p10-p11).*

2-Binding of Gal4 to the UASmut.
The authors make a lot of nice experiments in Figure 1 and 2 to decrease the dwell time of Gal4 to DNA and measure the consequences on gene activity. It is therefore a pity that the live-cell measurement of Gal4 binding is made on the Gal10 gene and not on Gal3 (Figure 3). To prove their point that the dwell time of Gal4 indeed controls burst duration, I encourage the authors to perform the live cell measurements of Gal4 binding to chromatin on Gal3 and to compare the WT and mutant UAS. This would make their point much stronger. The authors allude that this is technically challenging. In this respect, a system that might useful is the one described in PMID: 19651897, where a small gene array is used next to a LacO array, and allow to significantly amplify the signal of Gal4 binding.

We agree with the reviewer that measuring the binding of Gal4 to the Gal3 UAS variants would be very interesting. The current setup uses the PP7 transcription signal to locate the transcription site. As a consequence, the position of the locus can only be determined when transcriptionally active. Because of its high transcriptional activity, GAL10 TS is almost always present with very short off periods, allowing for colocalization with the Gal4 molecules. We have tried applying the same strategy to GAL3, so we can measure effect of the wildtype and mutated promoter. However, GAL3 transcription is much more infrequent, with long off periods between the on periods (average time

between bursts is ~4 min, Fig 1). Colocalization with Gal4 would rely on the position the TS of the previous or next transcriptional burst, and since there is significant movement of the TS between frames, the colocalizations are very unreliable. This is the main reason we performed this analysis on GAL10 and not on GAL3. Although the proposed gene array is an elegant system, it is questionable whether it is possible to use it reliably in our orbital tracking setup, because of its large size (larger than a point spread function). Therefore, this experiments with our current methodologies is not feasible and will require significant advances in the imaging systems to be successful. We have added a few sentences to better explain why we choose GAL3 for the mutational analysis and GAL10 for the Gal4 colocalization experiments (p5 and p8).

3-Figure 6.
I am not convinced by the measurements of burst duration presented here. Indeed, as seen in Figure 6A and as the authors state, the "bursts" (in the sense of the ON periods of the promoter) overlap, and the measurement of the burst size become somehow arbitrary. I am also not convinced by the autocorrelation analysis. To my knowledge, there is currently no published methods for how to fit an MS2/PP7 autocorrelation curve in the case of a two-state promoter model, and the assignment of the "burst" duration thus also appears not to be on very solid grounds (see for instance fig 5D where the two linear fits are not very good). Furthermore, the dwell-time of the nascent RNA appears to be in the same order as the Gal4 dwell time, and as the delay between bursts, complicating the situation. I am also worried that other processes might be at play that make the analysis of the correlation even less reliable. Are the autocorrelation curves invariant with respect to the duration of the acquisition?

The reviewer is correct that a full two-state autocorrelation function has not been published. Previous work (Larson et al. 2011, Lenstra et al, 2015) have used a linear fit to determine the dwell time of the RNA or dwell time of the burst (which we refer to as burst duration). However, even a full two-state theoretical formulation would not change the fact that the dwell time of Gal4, the burst separation, and the burst duration are on similar time scales, potentially complicating any fitting effort. For this reason, we have chosen the geometric approach we demonstrated and validated previously.

The average autocorrelation is calculated by averaging the individual autocorrelation curves using the arithmetic average method (averages the covariance functions, then normalize), which means that the autocorrelation curves are weighted for acquisition duration, similar to Lenstra et al. 2015, and are invariant with respect to the duration (See Coulon and Larson, 2016 for a complete description).

Finally, when the gene is fully induced, all the four binding sites of the Gal10 promoter are occupied, presumably by active Gal4 molecules, and is well possibly that bursting then becomes controlled by other factors, such as the lifetime of the PIC (which may be independent of Gal4 dynamics). One way to determine this would be to repeat the experiment with the UASmut and see whether it changes anything, both in terms of mature mRNA produced and dynamics. This figure may also be simplified or dropped if the authors wish to, since at present it does not convey essential conclusions.

The reviewer is correct that GAL10 contains four binding sites, but they are not all continuously occupied during full induction, as indicated by the fluctuations in the Gal4 channel in the orbital tracking experiments (performed at full induction conditions). The positive cross-correlation between Gal4 and GAL10 signals shows that these Gal4 fluctuations are directly coupled to GAL10 transcription. Because of the 4 binding sites at GAL10, the reviewers' suggestion of mutating the different UAS sequences at GAL10 would not be easy to interpret, which is why we specifically performed the UASmut experiments at GAL3. We have added a few sentences to better explain why we choose GAL3 for the mutational analysis and GAL10 for the Gal4 colocalization experiments (p5 and p8). The stability of other PIC components may also contribute to GAL10 transcription dynamics, but even if this is the case, our observation that Gal4 and GAL10 fluctuations are coupled indicates a large contribution of Gal4 dynamics in regulating transcription dynamics. More importantly, the primary message of Figure 6 is that the differential interaction of Gal4 with upstream signaling components such as Gal80 in the different galactose concentrations does not result in changes in burst duration, and therefore does not appear to influence Gal4 dwell time and/or the stability of the PIC. In order to more clearly explain the rationale behind the experiments

in Figure 6, we modified the explanation and incorporated a description of the role of the Gal80-Gal4 binding dynamics in regulating the burst frequency (p10-p11).

Minor comments:

-Figure 1: as a control, it is important to quantify the mRNA levels produce by these promoters to measure their relative strength.

We have now included the mRNA count per cell of the different promoters in EV1, determined by smFISH. The UASmut contains approximately 3 fold lower mRNA per cell than the UASwt.

-Figure 3: as a control, the authors should show that the Gal4-Halo is fully functional and does not lead to alteration of promoter regulation.

The functionality of Gal4-HALO was tested by growth on galactose and LiCl2 plates (Figure EV4A). On galactose-containing plates, Gal4-HALO cells showed similar growth as wildtype, while a Gal4 deletion strain does not show growth. In addition, growth of yeast on LiCl2 containing plates would indicates a loss of function, and unlike the gal4-deletion strain, a strain contaiing Gal4-HALO only shows minimal growth (although more than wildtype) (Figure EV4A). This indicates that Gal4-HALO is almost fully functional. We have expanded the description of this data in the manuscript to: " Addition of the HALO tag minimally affected its function, as Gal4-HALO cells showed similar growth as wildtype on galactose-containing plates and minimal growth on LiCl2 plates (Figure EV4A)" (p8).

-Figure 4F: can the measurements be done both in galactose and raffinose ?

In raffinose GAL10 is not expressed, and since this measurement relies on colocalization of Gal4 binding with GAL10 TS, with the current setup it is impossible to perform this measurement in raffinose as well.

-The authors could indicate on each panel the sugar conditions used, to facilitate reading.

We thank the reviewer for the suggestion, and have included the sugar conditions in the figures.

-In Figure 7, the authors may want to include Gal80 in their model.

We thank the reviewer for the suggestion and have adjusted the model of Figure 7 to include regulation by Gal80.

-For the orbital tracking, I was not sure of what happens when the transcription site goes off. Does the system keeps the same Z position waiting ?

The reviewer is correct, the systems waits in the same position until the TS appears again. For GAL10, the TS off time is very short. We have included a sentence in the material and methods section to clarify this.

Referee #2:

This paper uses some advanced microscopy to visualize events at a single gene locus in vivo. Monitoring both Gal4 activator and the transcript, the authors reach several conclusions. Not too surprisingly, Gal4 affinity correlates with dwell time in vivo and in vitro. The presence of nucleosomes decreases Gal4 dwell time (see comment below). Finally, looking at Gal4 binding versus the transcription burst duration and frequency, it appears that transcription is limited to the window when activator bound. Therefore, weaker Gal4 binding not only reduces Gal4 dwell time but also burst duration. In contrast, increased gene expression by galactose induction increases burst frequency, but not duration. The work is convincing and the observations provide interesting new insights into how we should think about transcription activation mechanisms.

I have some specific comments, which are mostly things that need some clarification:

p5. It would be helpful to know if "one GAL4 UAS" in the Gal3 promoter means it contains a single Gal4 binding site, or it's one UAS made up of multiple Gal4 binding sites.

We apologize for the confusion, we mean a single Gal4 binding site. We have adjusted this in the text to a single binding site (p5).

p5 The authors apply a threshold to the PP7 signals in Figure 1, the details of which are described in the methods section. As this seems to be based on the standard deviation in the background, is there some way to be sure this doesn't filter out single transcription events?

With the PP7-GAL3-MS2 construct we have independently determined that the elongation time of a single transcript is 30s (Figure 1G, p5). While we can't rule out that some individual events are below our threshold, the fact that we detect many events of 30s duration in the histogram confirms that most single transcription events are detected. We have included a reference in the main text to the material and methods section (p5) and added the following text to methods:"To determine the on and off periods, a threshold was applied to background subtracted traces of 2.5-3 times the standard deviations of the background. This number was chosen to reliably distinguish on and off periods from background levels at the single transcript level."

p6. If the UASmut only makes one transcript at a time, can that still be called a "burst"? This is a question of semantics, but producing single transcripts is also used as a distinguishing property of "constitutive" genes, so that definition becomes less useful if there can be single transcript "bursts".

We agree that these definitions are indeed confusing and have changed the text to: "The UAS mutation thus results in transcription of single RNAs like a constitutively transcribed gene and results in loss of bursting at the locus."

p6. I don't think you want to say that "competitor DNA was more effective in removing Gal4 from the labeled DNA". This is an equilibrium assay, so the competitor simply binds Gal4 that dissociates from the labeled DNA. The way it's written makes it sound like there's an active removal process.

We have adjusted the text to: "competitor DNA was more effective in competing with the labeled DNA for Gal4 binding" (p6).

Fig 2D - I recommend adding the red fluorophore into the schematics at the bottom (similar to 2A). Based on the picture, this must be in the middle of the probe?

The red fluorophore is not in the DNA, but in the histone as indicated in the schematics. We have adjusted the legends to clarify this.

MNase expt (Fig 3). Defining 95-225 nt fragments as nucleosomal size seems overly broad, especially since there's still debate about calling this type of weaker MNase protections as "fragile nucleosomes". Are the results any different if you set the bottom limit should be about 140-150? Other protein complexes can give footprints below that length. How can we be sure there's not something else there, perhaps even Gal4 plus other activators and coactivators?

We have rerun the analysis with higher thresholds (140-225 bp), which give the same results for all loci. As an example, we have included a plot for GAL3 with these higher thresholds in Figure EV3E. The Henikoff lab has recently shown that fragile nucleosomes are RSC-bound partially unwrapped nucleosomal intermediated (Brahma & Henikoff, 2019), suggesting that the fragments we detect at the GAL promoters are partially unwrapped nucleosomes. We have added this reference to the manuscript. Because partially unwrapped nucleosomes may have smaller footprints than 147 bp, we kept the lower bound threshold of 95 for the other figures.

p10 There's a sentence about the 152s duration of Gal10 transcription being compatible with a previously reported dwell time. It's unclear: dwell time of what? If this refers to the transcript appearance, that's a little confusing because this paper usually uses "on time" or "duration" to

describe that time period, while "dwell time" is usually used for Gal4 binding. The authors may want to check throughout the paper to make sure they are using terms consistently so readers don't get confused.

We thank the reviewer for the suggestion. We have adjusted the text on p10 to burst duration to make the nomenclature more consistent.

p10 I don't follow the logic of this conclusion:
"overlap between Gal4 binding and GAL10 RNA appearance agrees with our model that several RNA polymerases initiate while Gal4 is bound to the promoter"
Why couldn't there also be overlap for single transcription events? I think this point needs some further explanation to make the argument clear.

The reviewer is correct that the overlap in Gal4 and GAL10 signal does not exclude that there may also be overlap for single transcription events. We have removed this sentence from the manuscript.

Referee #3:

For a variety of technical reasons, linking the kinetics of transcription factor binding to the transcriptional output from a gene locus through real-time observation using optical microscopy has proved difficult. However, such knowledge is of fundamental importance in dissecting the mechanisms linking the two process. Multiple recent reports of transcriptional system with transient bursting behaviors that are not at steady-state highlight the importance of overcoming this challenge. Benjamin et al. have presented a comprehensive study on how the binding kinetics of Gal4 impact the subsequent transcriptional output of genes under its regulation. They combined in vitro and in vivo imaging approaches, supplemented with established genomics techniques where appropriate to establish a broader context in vivo, to characterize how different aspects of the GAL3 and GAL10 systems impact the kinetics of Gal4 binding, and, subsequently, how they change the behavior of transcription. Their observations lead to important conclusions, especially that the kinetics of Gal4 binding in their system is sufficient to produce bursting behavior without necessitating additional higher-order organizations in the nucleus. Their discovery of an upper-limit of Gal4 binding affinity for improving transcriptional output in vivo that is still significantly weaker than the consensus binding sequence for Gal4 is likewise intriguing, indicating that there are more parameters that constrain the evolution of the system beyond binding kinetics.

Overall, the experimental design is well thought out and the execution is also of good quality. Control experiments are done to limit artefacts when correlating signals and the limitations of each technique are also clearly explained and accounted for. While the experiments are conducted in yeast cells, the results and conclusions are sufficiently fundamental that they improve our basic understanding of transcriptional bursting and should remain applicable in more complex systems. If the authors could address the following questions/issues in the manuscript, I would recommend it for publication.

1. The authors stated the they chose the GAL10 system for the experiments in the second half of the manuscript but used the GAL3 system for the first half. Is there a reason they specifically chose GAL3 for the experiments presented Figures 1 and 2? Does GAL10 contain other elements that could complicate kinetic analysis?

The reason we chose GAL3 for the mutational analysis in the first half of the manuscript is because the GAL3 promoter has a single Gal4 binding site driving transcription, whereas the GAL10 promoter has 4 Gal4 binding sites. The spacing, configuration and cooperativity of the 4 UAS sites in the GAL10 promoter could contribute to transcriptional activation and bursting, which makes a mutational approach difficult to interpret. We therefore chose to use GAL3 to measure changes in bursting upon UAS mutations. We have added several sentence to the text to clarify this:
" Several often-studied galactose responsive genes, such as GAL1 and GAL10, contain multiple UASs (upstream activating sequences), of which the spacing, configuration and cooperativity may

contribute to transcription. Therefore, to study the role of TF binding in bursting, we focused on the GAL3 gene, which contains a single Gal4 UAS in its promoter. " (p5).
and
" For this experiment we used GAL10 instead of GAL3 because the GAL10 promoter contains four Gal4 UASs (instead of a single UAS for GAL3), which results in much higher transcriptional activity. Three GAL10 UASs are near consensus, the fourth one is the UASmut sequence we used for GAL3. We reasoned that this may increase the chance of observing colocalization of TF binding at its target gene with respect to nascent RNA production. " (p8)

If not, could they compare the two systems when their kinetics are measured in the manner of Figure 1 panels E, F, and I?
In this manuscript, we have measured both GAL3 (Fig 1) and GAL10 (Fig 6) transcription dynamics in full induction condition. Because the frequency of bursting is much higher for GAL10 than for GAL3, the on and off time as measured in Figure 1 panel E and F for GAL3 are difficult to perform for GAL10, because GAL10 bursts overlap as explained in figure 6. The GAL10 autocorrelation provides a burst duration is ~100s, which is larger than the 50s burst duration at GAL3 (fig 1E). From the traces it is clear that burst frequency is much higher for GAL10 than for GAL3, but the autocorrelation does not provide an absolute burst frequency that can be compared to the burst frequency of GAL3. Although the numbers are included in the text, we did not explicitly compare the kinetics in the text, because of the different analysis performed on the traces. To aid comparison of the two genes, we have added smFISH distributions of the number of nascent RNAs at the GAL10 and GAL3 TS (Figure EV5A and D)

Additionally of interest: Do the authors have information on where GAL10 wt would fall in Figure 2K compared to GAL3 wt?

The individual GAL10 binding sites (3 out of 4) have an affinity of near consensus, and are expected to fall near the UASconsensus in Figure 2K. The fourth GAL10 binding is the UASmut sequence. We have added this information to the manuscript (p8).
"Three GAL10 UASs are near consensus, the fourth one is the UASmut sequence we used for GAL3." (p8)

2. On page 8, line 19: the authors stated that they "observed single Gal4 molecules in the nucleus." Could the authors explicitly state their evidence for such a claim? Does the intensity over time change in discrete steps?

We observe single-step bleaching of the spots, indicating that they are indeed single molecules. There also exist some non-discrete intensity fluctuations that we attribute to the molecules moving in z, and therefore being irradiated by different amounts of laser light. Furthermore, our labeling density is such that very few (less than ~10) molecules are labeled in each nucleus, making it unlikely to have multiple labeled molecules in any one spot. We have included a sentence on the labeling density to the material and methods section: "Our labeling density is such that very few (less than ~10) molecules are labeled in each nucleus, making it unlikely to have multiple labeled molecules in any one spot."

3. Page 8, lines 30-32 and Figure EV4 B and C: The authors stated that H3 does not exhibit short-binding population. In Figure EV4 panels B and C, the authors showed a single-exponential fit for the H3 population. However, there seems to be a tail at longer time (>30~50 s) in the distribution that is not accounted for by the fit. If the H3 population is fit using a bi-exponential distribution, is there a significant population that is even longer lived? Is the "fast component" of H3 more stable than the "slow component" of Gal4? The captions in panel D suggests that the author did such a fit with H3.

We thank the reviewer for picking this up, the caption in panel D contained a typo and we have adjusted it. In addition, we have added a bi-exponential fit for the H3 data (at 200ms time interval), resulting in a large H3 population (60%) of molecules with a fast residence time of 7.2 +/- 0.3 s and a small population (5.4%) with a much slower residence time of 107.9 +/- 39.0 s, which does account better for the tail. Although we detect a slow population of H3 that is very long lived, this population represents only 5.4% of the H3 molecules. The majority of H3 molecules have a fast residence time of H3 (7.2s) that is dominated by bleaching and that is similar or less stable than the

slow residence time of Gal4 (6.7s ± 0.9s in raf and 17.1s ± 0.8s in gal). We therefore performed the same measurements with longer time intervals (1s). For easier comparison, we have included both the 1-exp and the 2-exp fit for H3, as well as the population percentage in Fig EV4B and D.

4. Page 8 the bottom half of the last paragraph, Figure 4 and EV4: What are the relative populations of the fast and slow components? Do the relative populations change between galactose and raffinose or does only the slow dwell time change?

We have included a figure with the relative populations of the fast and slow components at 200ms per frame in raffinose and galactose (Figure EV4E). The relative populations do not change between raffinose and galactose: the fast binding population is 25% in raf versus 29% in gal, the slow binding population is 6% in raf versus 5.6% in gal, and the unbound population is 69% in raf versus 65% in gal. We have also added the slow percentages to Figure EV4D. Thus, only the dwell time appears to change. We have included a sentence in the main text (p8).

5. Page 9, the paragraph spanning lines 12-23: The authors stated that H3 binds on much longer time scales and they used it as a control of the maximum measurable dwell time. However, for the 200 ms tracking experiments, H3 in Figure EV4D has a slow component that lasts for a shorter time than Gal4 in the presence of galactose. What is the reason for this discrepancy?

Figure EV4D contained an error in the axis labeling. The reason for this discrepancy is that the H3 data was fitted with a single exponential curve (reflecting the entire population), and Gal4 with a double exponent curve (reflecting only the long bound population). Using a bi-exponential fit for H3 results in a fast H3 residence time of 7s and a slow H3 residence time of 108s. We have now included both 1-exp and 2-exp fits for H3 in Figure EV4D. As was explained in the manuscript, to confirm that the Gal4 dwell time is not dominated by bleaching at 200ms, we repeated the experiment at 1000 ms. Longer interval does not increase Gal4 dwell time, but does result in an overall longer H3 dwell time.

6. Page 11 lines 12-14: Have the authors observed sufficient differences in photobleaching rates between orbital-tracking and time-lapse experiments that led to their suggestion that it could be a reason for the difference in the measured burst durations?

The imaging interval (262 ms vs 30s), laser powers and excitation mode (orbital confocal vs wide-field) are very different between the orbital tracking and time-lapse experiments, which may potentially explain the difference in burst duration. We included another potential explanation in the text (differences in signal to background ratios) as well as an explanation of how they may impact the burst duration: "The reasons for this discrepancy are not known at present but may be due to slightly different bleaching rates and/or signal-to-background ratios between the orbital tracking and time-lapse imaging methods. Bleaching or differential background during the acquisition can result in fluctuations in the data at longer time scales, which contribute to the slower decaying linear component of the autocorrelation, and may influence the determination of the burst duration." (p11)

7. In general about the discussion and Figure 7: The authors provided a very simplified figure as a summary of their findings and for possible mechanisms that they hypothesized based on their results. The authors should expand the figure such that it provides a more comprehensive summary of the salient properties of Gal4 driven regulation as explored in this manuscript and their hypothesized mechanism of action. For example, multiple RNA PolII recruited when TF is bound, termination of burst upod TF dissociation, how TF binding and burst kinetics are affected by various factors (binding site affinity, nucleosomes, galactose), etc.

We have expanded the figure to include more details of the regulation in the different conditions. The figure now shows that affinity and nucleosomes determine Gal4 dwell time and burst size. In low galactose, Gal80 repression of Gal4 results in lower burst frequency. In raffinose, Gal4 still binds DNA, but transcription is fully repressed by Gal80.

Minor issues:
1. Page 8, lines 21-23: The authors should cite the methods section describing how H3 movement is

used to select a threshold between frames used to determine if a particular Gal4 track is bound. It was unclear how this was done and what the threshold is when I first read this section and it was unclear where I could find the answer.

We have added the threshold and a reference to the methods section to the text.

2. Page 9 line 19: "Figures S4B-D" should be "Figure EV4B-D"

We have adjusted this.

2nd Editorial Decision 5th Apr 2019

Thank you for submitting your revised manuscript to The EMBO Journal. I have taken over its handling from my colleague Anne Nielsen, who has recently left our journal. Your manuscript has now been re-reviewed by the original referees, and I am pleased to inform you that all of them consider the study significantly improved and the earlier concerns adequately addressed. As you will see, only referee 1 still requests a minor modification, which I would ask you to incorporate during a final round of minor revision.

REFEREE REPORTS:

Referee #1:

In this revised version, the authors have answered and clarified several important points and I support publication of the manuscript.

One last point.
For autocorrelation curves of Fig 5D and 6C and the assignment of burst duration and frequency from these curves, the authors should add a word of caution because in particular for 5D, the double linear fit in 5D seems poor and the measurements are thus quite imprecise. For the figure 6C/D/E, I suggest that the authors add in the supplemental all the autocorrelation curves and the double linear fits. This will give the readers a more direct appreciation of the burst measurments.

Referee #2:

The authors have addressed my previous comments and I believe the paper is now acceptable for publication in the EMBO Journal.

Referee #3:

In this revision, the authors have addressed my questions put forward in my review of the original manuscript. Thus, I recommend that this manuscript be published.

Specifically, the information they added made their rationale for their choice of GAL10 vs GAL3 clearer and provided a comparison for the behavior between the two systems. Additionally, their clarification on their analysis of the binding kinetics answered my questions regarding them. The new Figure 7 is also much improved, providing a more comprehensive summary of their findings in graphical form.

EMBO PRESS

YOU MUST COMPLETE ALL CELLS WITH A PINK BACKGROUND ↓

PLEASE NOTE THAT THIS CHECKLIST WILL BE PUBLISHED ALONGSIDE YOUR PAPER

Corresponding Author Name: Tineke Lenstra
Journal Submitted to: EMBO Journal
Manuscript Number: EMBOJ-2018-100809

Reporting Checklist For Life Sciences Articles (Rev. June 2017)

This checklist is used to ensure good reporting standards and to improve the reproducibility of published results. These guidelines are consistent with the Principles and Guidelines for Reporting Preclinical Research issued by the NIH in 2014. Please follow the journal's authorship guidelines in preparing your manuscript.

A- 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.
- figure panels include only data points, measurements or observations that can be compared to each other in a scientifically meaningful way.
- graphs 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 and any statistical test employed should be justified
- Source Data should be included to report the data underlying graphs. Please follow the guidelines set out in the author ship guidelines on Data Presentation.

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 χ2 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.

Any descriptions too long for the figure legend should be included in the methods section and/or with the source data.

In the pink boxes below, please ensure that the answers to the following questions are reported in the manuscript itself. Every question should be answered. If the question is not relevant to your research, please write NA (non applicable). We encourage you to include a specific subsection in the methods section for statistics, reagents, animal models and human subjects.

B- Statistics and general methods

Please fill out these boxes ↓ (Do not worry if you cannot see all your text once you press return)

1.a. How was the sample size chosen to ensure adequate power to detect a pre-specified effect size?	NA
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	NA
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	For live-cell transcription (Figures 1 and 6, EV1), parts of traces were excluded if (1) the transcription site moves out of the field of view, (2) there was cell division during the tracking, or (3) the cell did not show a clear TS. For smFISH (Figures 1, 2, EV1, EV5), experiments were excluded if less than 500 cells were analyzed. For orbital tracking (Figure 4), sections of the fluorescence intensity traces were selected for active transcription by the appearance of signal above background in the carpet plots. The criteria for inclusion were pre-established. For all other experiments, all data was included.
3. Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. randomization procedure)? If yes, please describe.	NA
For animal studies, include a statement about randomization even if no randomization was used.	NA
4.a. Were any steps taken to minimize the effects of subjective bias during group allocation or/and when assessing results (e.g. blinding of the investigator)? If yes please describe.	NA
4.b. For animal studies, include a statement about blinding even if no blinding was done	NA
5. For every figure, are statistical tests justified as appropriate?	NA
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	When possible, the underlying distributions are shown either as main figure (Figure 1, 4, or as supplemntal data (Figure EV1, EV2, EV4, EV5), as well means wih error bars (SEM) of several replicates (Figure 1, 2, EV5). Since the differences are very clear with small error bars, and since the data was often not normally distributed, we did not apply a statistical test.
Is there an estimate of variation within each group of data?	The variation is shown in the error bars and is generally low. In case the variation was considerable (such as in figure 6), we used alternative methods to verify the same data (in Figure 6 we performed dose response in single cell).

USEFUL LINKS FOR COMPLETING THIS FORM

http://www.antibodypedia.com
http://1degreebio.org
http://www.equator-network.org/reporting-guidelines/improving-bioscience-research-repo

http://grants.nih.gov/grants/olaw/olaw.htm
http://www.mrc.ac.uk/Ourresearch/Ethicsresearchguidance/Useofanimals/index.htm
http://ClinicalTrials.gov
http://www.consort-statement.org
http://www.consort-statement.org/checklists/view/32-consort/66-title

http://www.equator-network.org/reporting-guidelines/reporting-recommendations-for-tun

http://datadryad.org

http://figshare.com

http://www.ncbi.nlm.nih.gov/gap

http://www.ebi.ac.uk/ega

http://biomodels.net/

http://biomodels.net/miriam/
http://jjj.biochem.sun.ac.za
http://oba.od.nih.gov/biosecurity/biosecurity_documents.html
http://www.selectagents.gov/

Is the variance similar between the groups that are being statistically compared?	Yes

C- Reagents

6. To show that antibodies were profiled for use in the system under study (assay and species), provide a citation, catalog number and/or clone number, supplementary information or reference to an antibody validation profile. e.g., Antibodypedia (see link list at top right), 1DegreeBio (see link list at top right).	NA
7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.	Sources of the cells are shown in EV Table 1

* for all hyperlinks, please see the table at the top right of the document

D- Animal Models

8. Report species, strain, gender, age of animals and genetic modification status where applicable. Please detail housing and husbandry conditions and the source of animals.	NA
9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.	NA
10. We recommend consulting the ARRIVE guidelines (see link list at top right) (PLoS Biol. 8(6), e1000412, 2010) to ensure that other relevant aspects of animal studies are adequately reported. See author guidelines, under 'Reporting Guidelines'. See also: NIH (see link list at top right) and MRC (see link list at top right) recommendations. Please confirm compliance.	NA

E- Human Subjects

11. Identify the committee(s) approving the study protocol.	NA
12. 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.	NA
13. For publication of patient photos, include a statement confirming that consent to publish was obtained.	NA
14. Report any restrictions on the availability (and/or on the use) of human data or samples.	NA
15. Report the clinical trial registration number (at ClinicalTrials.gov or equivalent), where applicable.	NA
16. 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.	NA
17. 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.	NA

F- Data Accessibility

18: Provide a "Data Availability" section at the end of the Materials & Methods, listing the accession codes for data generated in this study and deposited in a public database (e.g. RNA-Seq data: Gene Expression Omnibus GSE39462, Proteomics data: PRIDE PXD000208 etc.) Please refer to our author guidelines for 'Data Deposition'. Data deposition in a public repository is mandatory for: a. Protein, DNA and RNA sequences b. Macromolecular structures c. Crystallographic data for small molecules d. Functional genomics data e. Proteomics and molecular interactions	MNase data is deposited in GEO with accession number: GSE116337, reviewer token: chkfsocehnszjef. All microscopy data is available upon request.
19. Deposition is strongly recommended for any datasets that are central and integral to the study; please consider the journal's data policy. If no structured public repository exists for a given data type, we encourage the provision of datasets in the manuscript as a Supplementary Document (see author guidelines under 'Expanded View' or in unstructured repositories such as Dryad (see link list at top right) or Figshare (see link list at top right).	No public repository exists for microscopy data. Since the datasets are large, the images are stored on an internal server according to the FAIR principle. All data is available upon request.
20. Access to human clinical and genomic datasets should be provided with as few restrictions as possible while respecting ethical obligations to the patients and relevant medical and legal issues. If practically possible and compatible with the individual consent agreement used in the study, such data should be deposited in one of the major public access-controlled repositories such as dbGAP (see link list at top right) or EGA (see link list at top right).	NA
21. Computational models that are central and integral to a study should be shared without restrictions and provided in a machine-readable form. The relevant accession numbers or links should be provided. When possible, standardized format (SBML, CellML) should be used instead of scripts (e.g. MATLAB). Authors are strongly encouraged to follow the MIRIAM guidelines (see link list at top right) and deposit their model in a public database such as Biomodels (see link list at top right) or JWS Online (see link list at top right). If computer source code is provided with the paper, it should be deposited in a public repository or included in supplementary information.	Computational analysis were performed with previously available software (de Jonge et al, 2017, Mazza et al, 2013) or with custom scripts, which are described in the Material and Methods secontion and are available upon request.

G- Dual use research of concern

22. Could your study fall under dual use research restrictions? Please check biosecurity documents (see link list at top right) and list of select agents and toxins (APHIS/CDC) (see link list at top right). According to our biosecurity guidelines, provide a statement only if it could.	NA