molecular systems biology

Barriers to Transmission of Transcriptional Noise in a c-fos c-jun Pathway

Khyati Shah and Sanjay Tyagi

Corresponding author: Sanjay Tyagi, New Jersey Medical School, Rutgers University

Review timeline:

Submission date:	21 December 2012
Editorial Decision:	22 February 2013
Revision received:	03 June 2013
Accepted:	08 August 2013

Editor: Thomas Lemberger

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 22 February 2013

Thank you again for submitting your work to Molecular Systems Biology. First of all, I would like to apologize for the delay in getting back to you. We have now heard back from the three referees who agreed to evaluate your manuscript. As you will see from the reports below, the referees find the topic of your study of potential interest. They raise, however, substantial concerns on your work, which should be convincingly addressed in a revision of this work. The recommendations provided by the reviewers are very clear in this regard and refer to the potential contribution of transcript/protein stability and jun phosphorylation in noise attenuation. Reviewer #3 also raises a technical point on the characterization of the response of the proximity ligation assay.

In view of the quantitative nature of this work, we would kindly ask you to provide the 'figure source data' files (see http://www.nature.com/msb/authors/index.html#a3.4.3) for figures that present key quantitative data. While there is no current standard format for this type of data, we would be grateful if you could provide the data so that others may re-analyze the data, re-derive the statistics and potentially build upon your results. See additional instruction at http://www.nature.com/msb/authors/index.html#a3.4.3 and http://www.nature.com/msb/authors/source-data.pdf

If you feel you can satisfactorily deal with these points and those listed by the referees, you may wish to submit a revised version of your manuscript. Please attach a covering letter giving details of the way in which you have handled each of the points raised by the referees. A revised manuscript will be once again subject to review and you probably understand that we can give you no guarantee at this stage that the eventual outcome will be favorable.

REFEREE REPORTS

Reviewer #1

Review of: Barriers to Transmission of Transcriptional Noise in a c-fos c-jun Pathway
By Khyati Shah and Sanjay Tyagi

This is a very interesting manuscript from Tyagi's lab dealing with the central issue in biology of how order arises out of noise. Gene expression is noisy- we know this- but this leaves a major outstanding problem for cells and organisms, which is how to function despite this background problem. The problem has gained in recognition as developmental biologists and disease biologists have finally started to acknowledge their systems are heterogeneous and that this variation can be both a positive and negative this for system function.

A battleground for the understanding of noise control is the analysis of cell signaling and transcriptional pathways. Previous papers dealing with this issue have concentrated on simple eukaryotes or prokaryotes, and have concluded that pathways tend to amplify noise in gene expression. The subject has yet to be tackled in more complex cell types.

Tyagi's lab use here the single molecule RNA FISH method they have pioneered, which has now found use in many labs around the world. They also introduce a neat assay to detect the proximity of proteins, which is used to measure the heterogeneity in transcription factors dimer formation between cells.

To summarize the major findings of the work, the steady state transcript number per cells, for 2 dimerizing transcription factors, is highly heterogeneous even in cells undergoing a strong signaling shock. This heterogeneity is unlikely to be explained by simple coin-toss choices to transcribe by a polymerase, because the variance is simply far too high. The variability is less obvious when the level of the protein dimer is measured and the variability is again apparent when downstream transcripts are measured.

The authors conclude that dimer formation buffers the natural heterogeneity in transcript number and that this means that noise is not transmitted to a great degree between successive phases of a transcriptional response, therefore implying the noise buffering in complex cells is different from the published work in simple models.

I think the work merits publication in a high profile journal, as it is likely to be of interest to the single cell field and more thoughtful cell and developmental biologists. It should be of great interest to the wider cell biology field, and this will most likely be a long term impact.

I have some questions which should be resolved before publication.
1. Is the reduced heterogeneity in dimer formation simply a reflection of the different stabilities of transcript and protein? The variable snapshot given by the RNA counts would perhaps be less so if the sum of RNA production was integrated over time? You have the numbers clearly stated in the manuscript. Would it be possible to simulate the effects of these on the protein or dimer variance? Is it dimer formation or simply protein levels that are less noisy?
2. How good is the literature in assessing whether the c-fos-jun targets are actually solely the targets of the dimer, and not some other combination of fos (which is essential) and other factors? I realize evidence has been very clearly stated already, but if there is a significant contribution from other processes, this could mask any amplification? One approach would be to look at an annotated promoter for other obvious TF sites. Comparing different mammalian species is a good way to reveal control elements as they tend to be conserved. I'm sure this data is already out there.
3. Do you have any idea how sensitive the PLA is? Is it possible it only operates at clusters of dimers? The noise with RNA or protein appears different, but the noise is measured by different techniques, with likely different thresholds.

Reviewer #2

Shah and Tyagi present cutting-edge experimental results that suggest that there are barriers to the transmission of noise in the mammalian gene regulatory circuits that controls AP-1 activity. The manuscript is generally well written and a pleasure to read. In the end, I wish that it would go a little further in identifying the underlying mechanisms that generate this barrier.

The experimental work is outstanding and pushes the envelope of the analysis of gene regulatory networks in single cells. Combining smFISH with PLA to quantitate the number of AP1 dimers is stunning. The data is presented nicely, and a number of nice controls are included. It is clear that there is less cell-to-cell variability in the number of AP1 dimers, than of cfos/cjun mRNAs, or of downstream target genes. This supports the conclusion that there is a barrier to the transmission of noise.

But the work then stops short in identifying the mechanism that may provide such a barrier. The Introduction mentions protein halflife as a noise buffer, and the results/discussion mention heterodimerization. Both may indeed play a role. Quantitating the amount (though not necessarily the number) of cjun and cfos protein molecules is readily feasible and may distinguish between these two mechanism.

In addition, there is another feature of this gene regulatory network that surprisingly is not mentioned at all: that active AP1 requires phosphorylation of cjun by JNK constituting a feedforward loop that may also have noise-attenuating properties. ((The abundance of p-jun is also readily quantifiable in single cells.) I am not sure whether the PLA results require this phosphorylation or not. If they don't then cell-to-cell variability of JNK may contribute to downstream target gene expression. Indeed, while the authors did a nice cross-correlation analysis of cjun and cfos mRNA data, it would also be of interest to show whether expression of the two downstream target genes are correlated or not.

Finally, it strikes that the present study and the generation of such nice quantitative data of various points in the gene regulatory circuit, would benefit tremendously from a computational modeling component that would allow quantitative conclusions about the various mechanisms that produce noise and dampen its transmission. However, in my opinion such a modeling component is beyond the scope of the present study; I do believe the experimental work (substantiated as suggested above) could stand on its own.

Reviewer #3

The work submitted by Shah and Tyagi consists of a study of the variability of various gene expression products induced by serum induction. Their experimental work, descriptive in nature, is elegant, carefully done and the data is solid. Using a combination of single molecule mRNA FISH and a proximity ligation assay, they record simultaneously the copy number of c-fos and c-jun mRNAs, the copy number of the complex their protein products form, and they also measure the copy numbers of downstream mRNAs induced by the complex. This exhaustive approach gives them the potential to identify how noise is transmitted through the signaling pathway. Surprisingly, they show that high levels of noise at the levels of the immediate responding genes mRNA are not transmitted by the protein complex they encode. Expression of the downstream mRNAs is in turn relatively insensitive to the protein complex concentration.
The results are novel and interesting. However, the study is restricted to a purely descriptive and somewhat superficial level that does not yield any definitive mechanistic answer. For instance, the most provocative result of the study is the fact that noisy, uncorrelated expression of c-fos and c-jun mRNA results in a relatively homogenous distribution of c-jun/c-fos protein dimer. It is quite disappointing that the authors did not address the mechanism of this effect. Does the mRNA noise become buffered at the level of the protein product of each gene, or is the dimerization the main correcting effect?
The main experimental issue I see with the present work is the contention from the authors that the PLA signal is proportional to the concentration of dimers. A single-cell control is necessary here to show that concentration of dimers in individual cells is indeed proportional to the number of observed PLA spots. This control is crucial for this study since both the average level of dimers and their variability are central to the results. A way to perform this control would be to add an exogenous epitope (e.g. Flag or Myc) to their artificial dimer construct, and perform both the PLA

and Immunofluorescence against the epitope. The number of PLA spots should be proportional to the immunofluorescence signal at the single cell level.

1st Revision - authors' response 03 June 2013

Reviewer 1.

This is a very interesting manuscript from Tyagi's lab dealing with the central issue in biology of how order arises out of noise. Gene expression is noisy- we know this- but this leaves a major outstanding problem for cells and organisms, which is how to function despite this background problem. The problem has gained in recognition as developmental biologists and disease biologists have finally started to acknowledge their systems are heterogeneous and that this variation can be both a positive and negative this for system function.

A battleground for the understanding of noise control is the analysis of cell signaling and transcriptional pathways. Previous papers dealing with this issue have concentrated on simple eukaryotes or prokaryotes, and have concluded that pathways tend to amplify noise in gene expression. The subject has yet to be tackled in more complex cell types.

Tyagi's lab use here the single molecule RNA FISH method they have pioneered, which has now found use in many labs around the world. They also introduce a neat assay to detect the proximity of proteins, which is used to measure the heterogeneity in transcription factors dimer formation between cells.

To summarize the major findings of the work, the steady state transcript number per cells, for 2 dimerizing transcription factors, is highly heterogeneous even in cells undergoing a strong signaling shock. This heterogeneity is unlikely to be explained by simple coin-toss choices to transcribe by a polymerase, because the variance is simply far too high. The variability is less obvious when the level of the protein dimer is measured and the variability is again apparent when downstream transcripts are measured.

The authors conclude that dimer formation buffers the natural heterogeneity in transcript number and that this means that noise is not transmitted to a great degree between successive phases of a transcriptional response, therefore implying the noise buffering in complex cells is different from the published work in simple models.

I think the work merits publication in a high profile journal, as it is likely to be of interest to the single cell field and more thoughtful cell and developmental biologists. It should be of great interest to the wider cell biology field, and this will most likely be a long term impact.

I have some questions which should be resolved before publication.

1. *Is the reduced heterogeneity in dimer formation simply a reflection of the different stabilities of transcript and protein? Is it dimer formation or simply protein levels that are less noisy?*

Our experimental evidence points to the stability as being the key factor. However, we now distinguish between the heterogeneity in individual proteins and in the heterodimers formed by them (Figures S1 &S2). New results show that the coefficient of variation in c-fos protein is similar to that of its own mRNA, whereas, the coefficient of variation in c-jun protein is lower then its own mRNA. The later is almost the same as the heterodimers. Additional factors that might play a role are discussed on pages 7,8,10 and 11.

2. *The variable snapshot given by the RNA counts would perhaps be less so if the sum of RNA production was integrated over time? You have the numbers clearly stated in the manuscript. Would it be possible to simulate the effects of these on the protein or dimer variance?*

That is definitely the case. We have shown earlier (Raj et al, 2006) that higher stability of proteins buffer against fluctuations in mRNA production. This was shown by simulation, as well by experiments in which we decreased the stability of a protein so that it would be about the same as that of mRNA (Raj et al, 2006, Figure 7). With this change, the buffering was dissipated and the heterogeneity in proteins began to mirror the heterogeneity in RNA. We presume increased stability of either, mRNAs, proteins, or, their complexes would have similar effects in the present context.

2. How good is the literature in assessing whether the c-fos-jun targets are actually solely the targets of the dimer, and not some other combination of fos (which is essential) and other factors? I realize evidence has been very clearly stated already, but if there is a significant contribution from other processes, this could mask any amplification? One approach would be to look at an annotated promoter for other obvious TF sites. Comparing different mammalian species is a good way to reveal control elements as they tend to be conserved. I'm sure this data is already out there.

This issue is central to our work - the entire edifice rests of the hypothesis that the expression of downstream genes strictly depends on the c-fos and c-jun and some other factor, also produced in response to serum, can't take their place. The available evidence for the involvement of c-fos and c-jun is particularly strong for collagenase 1 and carefully described on page 8. Among the key evidence is a chromatin immunoprecipitation study focused on collagenase 1 (Martens et al 2003). They show that c-fos and c-jun get attached to the collagenase 1 promoter after the addition of serum along with an accessory transcription factor p300, which is constitutively present in the cell (also known as E1A), and then recruit gene induction apparatus. A second key evidence is that cell lines in which c-fos is knocked out, do not produce collagenase 1 in response to serum (Hu et al 1994). A third piece of evidence, that addresses your concern that can some other factor take the place of c-jun, comes from the studies of Firestein (J Clin Investigation, 108, 73-,2001) and colleagues who studied the induction of collagenase 1 in synoviocytes. These cells express c-fos and c-jun, when stimulated by IL-1 and produce collagenase 1. If this induction is performed in presence of an inhibitor of c-jun phosphorylation (necessary for c-jun activity) collagenase 1 is not produced. So, both c-fos and c-jun are necessary.

We explored ENCODE data that has been assembled about genome wide transcription factor occupancy (Figure 1). However, this data is of limited utility because it is derived from HeLa cells growing under steady state, a condition under which collagenase 1 is not expressed. Nonetheless, we are attaching that data for your consideration. This shows small amounts of c-fos and c-jun on the collagenase 1 promoter (shaded). However, the data is likely to represent the off state of the gene (and in average cells) as Pol II is almost completely absent.

Comparison of transcription factor binding between species has so far been done for only two transcription factors, CEBPA and HNF4A (Schmidt, et al, Science, 328, 1036-, 2010). Please let us know if there is a study that would be more useful in this regard.

Nonetheless, the targeted studies that we site on page 8 provide strong basis of our analysis.

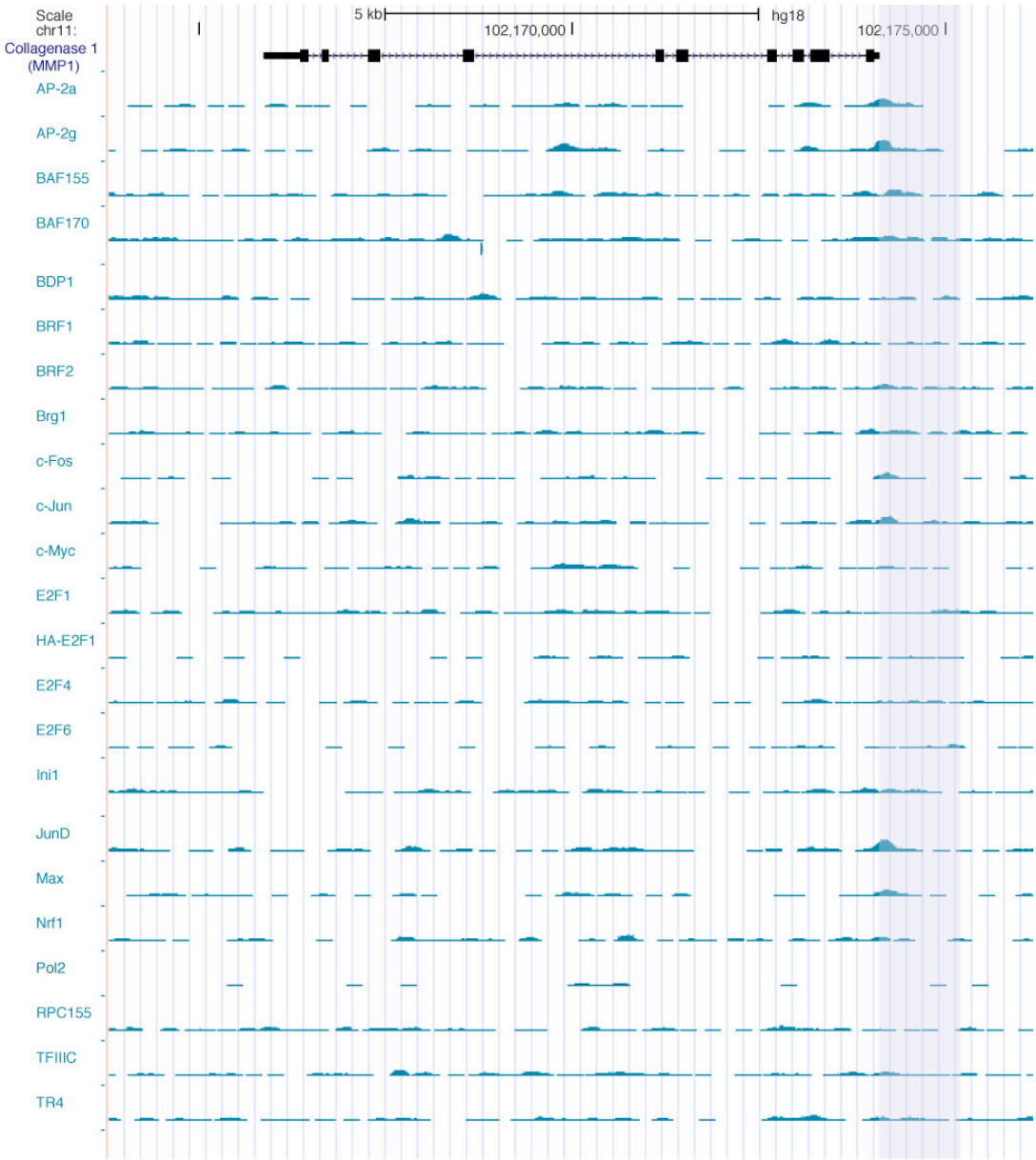

Figure 1. Occupancy of collagenase 1 gene by transcription factors and Pol II in HeLa cells. ENCODE data.

3. Do you have any idea how sensitive the PLA is? Is it possible it only operates at clusters of dimers? The noise with RNA or protein appears different, but the noise is measured by different techniques, with likely different thresholds.

We think PLA is highly specific but not very sensitive. Only a small fraction of hetrodimers are likely to produce a spot. However, as suggested by the reviewer 3, it is more important that the number of PLA signals be proportional to the actual dimmers present in the cell. Upon his suggestion we undertook a study to establish that proportionality relationship. The results are shown in Fig. 5 and are discussed on page 7.

If there was a tendency for the heterodimers to form clusters (or if they were to be differentially concentrated in some nuclear structure), we will indeed expect those clusters to yield PLA spots with higher probability. Direct immuno fluorescence images for c-fos and c-jun do not show any tendency of the two proteins to form colocalizing clusters (c-fos distribution in the nucleus is granular, whereas, the c-jun distribution is smooth) (Fig. S1). Once these proteins heterodimerize

through their leucine zipper domains, there is no other structure left in them that can potentially mediate a multimerization reaction.

Your point about noise measurements with techniques with difference threshold is valid. To address this we are also presenting the coefficients of variation in Fig. S2. That parameter is insensitive to the units of measurement. The results are qualitatively similar - the noise in heterodimers is lower than both upstream and downstream mRNAs.

Reviewer 2.

Shah and Tyagi present cutting-edge experimental results that suggest that there are barriers to the transmission of noise in the mammalian gene regulatory circuits that controls AP-1 activity. The manuscript is generally well written and a pleasure to read. In the end, I wish that it would go a little further in identifying the underlying mechanisms that generate this barrier.

The experimental work is outstanding and pushes the envelope of the analysis of gene regulatory networks in single cells. Combining smFISH with PLA to quantitate the number of AP1 dimers is stunning. The data is presented nicely, and a number of nice controls are included. It is clear that there is less cell-to-cell variability in the number of AP1 dimers, than of cfos/cjun mRNAs, or of downstream target genes. This supports the conclusion that there is a barrier to the transmission of noise.

But the work then stops short in identifying the mechanism that may provide such a barrier. The Introduction mentions protein halflife as a noise buffer, and the results/discussion mention heterodimerization. Both may indeed play a role. Quantitating the amount (though not necessarily the number) of cjun and cfos protein molecules is readily feasible and may distinguish between these two mechanism.

We also think they both play a role. We are now presenting measurements of cellular amounts of c-fos and c-jun proteins and their cell-to-cell variations (Figures S1&S2). c-fos protein displays about similar level of variation as its own mRNA, whereas, c-jun protein exhibits a variation that is less then its own mRNA and about the same as the heterodimers. This suggests that the stability of individual proteins contributes towards lowering the heterogeneity. Based on existing literature we hypothesize that the proteins get stabilized further upon heterodimerization. Please see discussion on pages 7,8,10 and 11.

In addition to this stability-based barrier, we postulate that chromatin itself serve as a barrier to the transmission of noise. Our data suggests (Fig. 6C) that the induction of downstream genes depends on the presence of heterodimers but the probability of induction does not increase with the number of heterdimers in the cell. An implication is that if the heterodimers were to be more variable, the variability in the downstream genes wouldn't be more variable. Please see discussion on pages 7,8,10 and 11.

In addition, there is another feature of this gene regulatory network that surprisingly is not mentioned at all: that active AP1 requires phosphorylation of cjun by JNK constituting a feedforward loop that may also have noise-attenuating properties. ((The abundance of p-jun is also readily quantifiable in single cells.) I am not sure whether the PLA results require this phosphorylation or not. If they don't then cell-to-cell variability of JNK may contribute to downstream target gene expression.

We apologize for the omission. One page 4 we added a sentence about phosphorylation of c-jun by JNKs being important for target gene activation and on pages 10 and 11 we present an hypothesis about its possible role in noise reduction. The antibody for c-jun that we use in PLA does not require it to be phosphorylated for detection. The antibody (Santa Cruz antibody SC-1694) is a polyclonal antibody raised against first 79 amino acids in the c-jun coding sequence. Although, this stretch includes amino acids 63 and 73 that are phosphorylated by JNK, the probability that only those two phosphorylated residues are the recognized epitopes is extremely low.

Your insight that c-jun phosphorylation may have noise attenuation properties is intriguing and the possibility that variation in JNK may control the variation in active c-jun is real. We have added this as one of the hypothetical noise reducing factors on pages 10 and 11. Since we are detecting heterodimeric complexes without regard to the phosphorylation of status of their c-jun component, we don't know what fraction is phosphorylated. Your point that downstream genes are responsive

to the phosphorylated form, rather than the unphosphorylated fraction is right, but it can't be addressed using our tools.

While the authors did a nice cross-correlation analysis of cjun and cfos mRNA data, it would also be of interest to show whether expression of the two downstream target genes are correlated or not.

We are now presenting cross-correlation analysis of the two downstream genes (Pages 9 and 10 and Fig. S3). Given the low probability of induction of each gene in single cells, there was little correlation between the two.

Finally, it strikes that the present study and the generation of such nice quantitative data of various points in the gene regulatory circuit, would benefit tremendously from a computational modeling component that would allow quantitative conclusions about the various mechanisms that produce noise and dampen its transmission. However, in my opinion such a modeling component is beyond the scope of the present study; I do believe the experimental work (substantiated as suggested above) could stand on its own.

We appreciate the need for the modeling but unfortunately lack the expertise. We will be happy to assist others who are more qualified to perform the modeling studies. Towards this goal, we are including a carefully selected set of data (cell-by-cell measurements of mRNAs, proteins, and heterodimers for upstream and downstream genes) (Table 1). Also, we will provide primary images and image analysis software for anybody who is interested.

Reviewer 3.

The work submitted by Shah and Tyagi consists of a study of the variability of various gene expression products induced by serum induction. Their experimental work, descriptive in nature, is elegant, carefully done and the data is solid. Using a combination of single molecule mRNA FISH and a proximity ligation assay, they record simultaneously the copy number of c-fos and c-jun mRNAs, the copy number of the complex their protein products form, and they also measure the copy numbers of downstream mRNAs induced by the complex. This exhaustive approach gives them the potential to identify how noise is transmitted through the signaling pathway. Surprisingly, they show that high levels of noise at the levels of the immediate responding genes mRNA are not transmitted by the protein complex they encode. Expression of the downstream mRNAs is in turn relatively insensitive to the protein complex concentration.
The results are novel and interesting. However, the study is restricted to a purely descriptive and somewhat superficial level that does not yield any definitive mechanistic answer. For instance, the most provocative result of the study is the fact that noisy, uncorrelated expression of c-fos and c-jun mRNA results in a relatively homogenous distribution of c-jun/c-fos protein dimer. It is quite disappointing that the authors did not address the mechanism of this effect. Does the mRNA noise become buffered at the level of the protein product of each gene, or is the dimerization the main correcting effect?

We are now presenting measurements of cellular amounts of c-fos and c-jun proteins and their cell-to-cell variations (Figures S1&S2). c-fos protein displays about similar level of variation as its own mRNA, whereas, c-jun protein exhibits a variation that is less then its own mRNA and about the same as the heterodimers. Please see discussions at pages 7,8,10 and 11 for the mechanisms responsible for the phenomenon.

The main experimental issue I see with the present work is the contention from the authors that the PLA signal is proportional to the concentration of dimers. A single-cell control is necessary here to show that concentration of dimers in individual cells is indeed proportional to the number of observed PLA spots. This control is crucial for this study since both the average level of dimers and their variability are central to the results. A way to perform this control would be to add an exogenous epitope (e.g. Flag or Myc) to their artificial dimer construct, and perform both the PLA and Immunofluorescence against the epitope. The number of PLA spots should be proportional to the immunofluorescence signal at the single cell level.

Very smart idea. Thank you for suggesting it. We made the construct as you suggested with the FLAG tag, integrated it within HeLa cell genome, and then isolated stable cell lines. We were able to show that the fusion protein yields PLA spots when probed by a mouse c-fos antibody and rabbit c-jun antibody and also yields FLAG specific signals when probed with a third goat antibody. The

data is presented in New Figure 5. Even though the correlation between FLAG signal and PLA spot was only 0.57, the proportionality is clearly apparent.

