Evaluation of Hansen et al.: Nuance Is Crucial in Comparisons of Noise

Noah Olsman,[1] Fangzhou Xiao,[2] and John Doyle[1,*]
[1]Department of Control and Dynamical Systems, California Institute of Technology, 1200 E. California Blvd, Pasadena, CA 91125, USA
[2]Division of Biology and Biological Engineering, California Institute of Technology, 1200 E. California Blvd, Pasadena, CA 91125, USA
*Correspondence: doyle@caltech.edu
https://doi.org/10.1016/j.cels.2018.10.003

One snapshot of the peer review process for "Cytoplasmic Amplification of Transcriptional Noise Generates Substantial Cell-to-Cell Variability" (Hansen et al., 2018).

Editor's Note: This is a first-round review of "Cytoplasmic Amplification of Transcriptional Noise Generates Substantial Cell-to-Cell Variability" by Leor Weinberger, Maike Hansen, and their colleagues; it was written for Cell Systems *as part of the peer review process. We chose to feature it here because its nuanced treatment of noise, Hansen et al. (2018, this issue of* Cell Systems*), and Battich et al. (2015) exemplifies scholarship. The constructive critique it presents also improved Hansen et al. (2018) without imposing an agenda on its authors. After the first round of review, Hansen et al. (2018) was revised to take the reviewers' comments into account, re-submitted, re-reviewed, accepted for publication, and then published in this issue of* Cell Systems*. For comparison, an earlier version of Hansen et al. was deposited on bioRxiv ahead of review and can be found here: https://doi.org/10.1101/222901. Olsman et al. chose to reveal their identities during the peer review process within this peer review. Hansen et al. support the publication of this Peer Review; their permission to use it was obtained after their paper was officially accepted. This Peer Review was not itself peer reviewed. It has been lightly edited for stylistic polish and clarity. No scientific content has been substantively altered.*

INTRODUCTION

The reviewers are generally enthusiastic about this paper but have many suggestions for how it could be improved. Thus, we hope any critical tone is not misinterpreted, and our assessment of "major revision" is probably harsh, but "minor" seemed too tepid. We are eager to help in any way we can and don't want to substantially delay this important paper. The official reviewer is John Doyle, but the hard work in this review was done largely by two of my graduate students, Noah Olsman and Fangzhou (Fang) Xiao. We are fairly theory oriented, though Noah and Fang know a lot of biology.

Overview and Summary

The paper by Hansen et al. investigates the nature of eukaryotic mRNA fluctuations using both simulation and experimental results. In particular, the authors seek to understand whether the specific process of nuclear export serves as a mechanism for noise attenuation, as implied by the work of Battich et al. (2015) (referred to below as simply "Battich et al.") when specific references therein are not being made), or if the export process amplifies noise. Hansen et al. (2018) conclude that in general, the latter is the better characterization of what they observe. They support this argument in several ways.

First, the authors argue that the Fano factor (rather than the coefficient of variation (CV) or CV^2, as used by Battich et al.) is a better tool for measuring relative changes in noise between Poisson-like processes because of its invariance to changes in mean values (where the CV of a Poisson process scale with the inverse square root of the mean). Next, the authors show via simulations that a chemical master equation model of mRNA synthesis consistently exhibits noise attenuation in a parameter regime that they say is representative of ~70% of genes. Finally, the authors provide detailed experimental results that use FISH measurements to quantify the distribution of mRNA for specific endogenous and viral genes in both the cytoplasm and the nucleus.

Because of the multifaceted approach the authors took in approaching the problem, we will break the review into three sections: first, we address the theoretical arguments regarding noise quantification; second, we will address the modeling and simulation results; and third, we will discuss the experimental work.

The overall conclusion of this review is that the authors' results are sound; however, some of the interpretations of these results are a bit too broad and more definitive than is justified. While the authors' interpretations are generally reasonable, there are enough systematic differences between their work and prior work (particularly between their paper and the Battich et al. paper) that it is difficult to say conclusively whether the results of each paper can be directly compared in a meaningful way. Specifically, the authors measure noise differently, use a different underlying mathematical model, and use different cell lines. It is not essential that all the issues we raise be addressed in this paper, but simply that this paper facilitate—as much as possible—their eventual resolution.

Theoretical Motivation for Use of Fano Factor, as Done by Hansen et al., versus CV or CV^2, as Done by Battich et al.

We think there is a subtle but important distinction between quantifying noise in some absolute sense and quantifying the relative effects of particular mechanisms on noise in a given application. In the former case, where we are comparing the variability of a process to a specific signal (e.g., the mean of the distribution), it makes sense to compare the size of fluctuations to the size of the signal. In this sense, CV is natural, though it is better

to look at both the mean and the CV (or variance) together. It is true that the CV scales with the mean, but scaling the mean can potentially be a perfectly reasonable approach to reducing noise in many situations. Indeed, much of engineering uses this or related high-gain feedback strategies to provide robustness at the expense of efficiency.

If, however, the desire is to compare the noise from a given mechanism against a baseline distribution in a statistically meaningful way, then CV has shortcomings for the reasons the authors describe. In particular, it makes sense to compare the measured distribution to an equivalent distribution with the same mean. In this case, one could use Fano factor to compare distributions. However, if the real goal is to compare distributions, then the Fano factor has issues because it is only a function the first two moments of the distribution. If the real goal is to compare distributions, then a test like the Kolmogorov-Smirnov test is a more direct measure.

The Fano factor has the desirable feature of both quantifying the magnitude of variations and quantifying distance from a Poisson distribution (a reasonable baseline for biological processes), but it does neither of these things as well as the other two metrics do individually. The purpose of this discussion is not to criticize the particular use of Fano factor in Hansen et al. but to encourage the authors to include a more informative discussion of the pros and cons of different noise metrics. This could be achieved by more explicitly discussing the limitations of the Fano factor in noise quantification. Fano factor definitely works to serve the purposes for which the authors use it in the paper, but it does come with tradeoffs that shouldn't be ignored, and it isn't objectively superior.

In general, we think that the authors could simply argue that if you want to have a convenient way to compare both magnitude and distributional properties of two different species (e.g., nuclear and cytoplasmic mRNA), Fano factor plots like those in Figures 1E and 2F in Hansen et al. (2018) are a concise means of doing this. If one were to do the same thing with, say, CV and mean mRNA or Kolmogorov-Smirnov statistics for each species, it would require a complicated 4D (or worse) plot. Fano factor yields a nice balance for this particular type of statistical question about comparing noise.

Simulation Results and Arguments that $FF_{nuc} < FF_{cyt}$

We think the actual simulation results in Hansen et al. are compelling and give reasonable evidence that we should expect the measure of noise amplification caused by nuclear export used by the authors to be larger than 1 for most physiological parameters, though we also think that they are not sufficient to directly contradict the results of Battich et al. (2015). This is because the Battich paper models nuclear export as a diffusive process, whereas the present authors model export as a single reaction rate constant. These modeling assumptions likely lead to different export statistics and may very well yield different results regarding noise amplification/attenuation regardless of the way in which noise is measured.

Hansen et al. should either argue why the simple reaction rate model of export is a better (or at least no worse) model of nuclear export than the diffusive model from Battich et al. or should make it clear that this modeling assumption could have a strong effect on the final results. In the latter case, the claims that, "the data constraints show that ~85% of genes fall in the parameter regime in which noise is amplified in the cytoplasm and only about 2.5% of genes fall in the parameter regime where noise is attenuated down to minimally stochastic Poisson levels—substantially less than previously implied (Battich et al., 2015)" (quoted from the originally submitted version of Hansen et al.) should be modified, as it is unclear if the conclusion is valid. The arguments about parameter regimes strictly depend on the underlying model, as such it is not clearly possible to argue if the fraction of genes amplifying versus attenuating noise is higher or lower than what is implied by Battich et al.

One way to ameliorate this difficulty would be to implement a simplified model of discrete diffusion in the current paper, for example having linearly arranged nuclear states.

Nascent nuclear mRNA $\rightarrow N_0 \leftrightarrow N_1 \leftrightarrow \ldots \leftrightarrow N_{final} \rightarrow$ Cytoplasmic mRNA $\rightarrow$ degradation

This might strike a nice balance between capturing the dynamics of diffusion and still allowing the simulations to scale well enough to cover a wide parameter range. While the reviewers are not particularly familiar with this area, the review paper "Dynamics of transcription and mRNA export" (Darzacq et al., 2005) claims that there is strong evidence for diffusion driving mRNA export. It may be that, for realistic parameters, the reaction rate model is good enough for the purposes of this paper. However, this should at least be probed with simulations, especially when it seems relatively easy to do.

The authors also provide some theoretical arguments about why we would expect cytoplasmic noise to be higher than nuclear, largely based on using approximations and bounds derived from noise bandwidth properties. While we don't have a reason to think these methods are fundamentally incorrect, we did find at least one (possibly fixable) mathematical error in the argument. In particular, in the statement that "for all cases $\frac{Noise_{cyt}}{Noise_{nuc}} = \frac{C}{N}\frac{\xi_{cyt}}{\xi_{nuc}}$ (Eqn. 1) reduces to $\frac{Noise_{cyt}}{Noise_{nuc}} \leq \frac{C}{N}$, which predicts that there is a strong tendency for $Noise_{cyt} > Noise_{nuc}$ when $C > N$" (quoted from the originally submitted version of Hansen et al.), the "tendency" argument doesn't logically follow. If $\frac{a}{b} \leq \frac{c}{d}$, and c > d, this does not imply that a > b. For example, 3/4 < 2/1, but 2 > 1 does not imply 3 > 4. The authors should be more careful drawing conclusions from inequalities. We didn't actually check the rest of the noise bandwidth arguments past this in much detail, but they seem potentially unnecessary if other more direct approaches are fleshed out more clearly.

For example, it might be easier to address these theoretical questions by directly computing statistics of the model at hand. Because the underlying system is a linear chemical master equation, it is possible to explicitly compute CV, Fano factor, etc. The reviewers performed this calculation, and it looks promising; however, the terms are quite complicated. We are happy to share this work if the authors find it useful. (There are also a variety of numerical methods that should be effective for these models.) While there were no obvious nice simplifications to be made, it may be possible to analyze the physiological ranges of interest to get explicit approximations of the model's statistics.

The main reason that this might be useful is that it is a bit confusing to add yet another measurement of variability (noise bandwidth) on top of the CV and Fano factor already discussed. It also may make it easier to tell precisely when we expect cytoplasmic variability to exceed nuclear.

(As a note, it is somewhat confusing why the authors provide a lengthy argument as to why FF is a better metric of biological noise but then in Figures 1F–1H define the nuclear/cytoplasmic noise ratio in terms of CV^2. Equation 1 seems to refer to this same ratio as being between Fano factors, is one of these descriptions a mistake, or are the two definitions equivalent for some reason that we missed? This should be clarified.)

Experiments and Statistical Analysis

As the reviewers are not experimentalists, we cannot really comment directly on those details of the paper. For the analysis of the data, a strong point of the current work is the direct comparison of nuclear and cytoplasmic noise for individual genes. Ideally, the authors would be able to use the same exact statistical model from Battich et al.; however, we appreciate that this may not be feasible. While the authors do incorporate some physiological parameters to adjust their statistics, these appear to be far less in-depth than the Battich et al. methods, and thus, it is hard to distinguish directly if the experimental data from the current paper actually disagrees with the phenomena observed in previous work, e.g., whether or not physiological variability accounts for most of the variation in cytoplasmic mRNA concentrations.

While we don't view this as a fundamental barrier to publication, it is unclear whether the extrinsic noise filtering performed in the current work is comparable to that of Battich et al. Given that the Battich methods appear to be extremely accurate at predicting mRNA distributions, weight should be given to their methodology. For example, regarding Figure 2E, Hansen et al. (2018) make the point that the data show Fano factor > 1, implying super-Poisson noise. However, the claim of minimal noise in the Battich paper is contingent on specifically measuring intrinsic noise, using their method of separating extrinsic variability from intrinsic, so even based on their results you would likely expect super-Poissonian noise (regardless of how it is quantified) from the unfiltered data.

At minimum, the authors should make it clear that while their data provides evidence for increased cytoplasmic noise, they cannot rule out that physiological variations account for this increased mRNA variability. This is especially true because Battich et al. (2015) uses different methods to filter extrinsic noise and uses different metrics for noise (CV, Kolmogorov-Smirnov fit to Poisson, prediction strength). The best result would be to use the methods from Battich et al. (2015) to process the current paper's data; this would make for an extremely convincing case that cytoplasmic mRNA noise is indeed amplified. We acknowledge that this would likely require processing technique that may not be easily implemented on the current paper's data. If the authors choose not to use these methods, then they should make clear in the paper why they believe the methods they have used are superior or explain the shortcomings of their approach.

Evolutionary Arguments

Another argument the authors used to motivate the research is that Fraser et al. (2004) showed evolutionary selection on promoter structure in Eukaryotes, suggesting that noise must be large enough to drive this selection, and therefore cannot be fully attenuated by a passive mechanism such as nucleus. We think this claim actually supports the point made by Battich et al. (2015). The buffering effects of nucleus exportation usually isn't powerful enough to reduce noise to a level very close to the Poisson level, as shown in Figure 6E, in Battich et al. (2015). If the Fano factor for the cytoplasmic data point were computed in Figure 6E, it could be easily seen that the Fano factors are not close to 1, but could range as high up as 20 or 30, depending on the mean. The noise is not very small in terms of CV^2 either, especially given how spread out the dots are. So, it could be the case that nuclear export attenuates noise, and was actually evolutionarily selected, as was argued briefly in line 14 of Page 1608 in Battich 2015. At the same time, because nuclear buffering was developed later, or because a long buffering time was disadvantageous, or because nuclear buffering does not attenuate noise enough, there could still be strong selection for promoter architecture to attenuate noise.

Also, this evolution argument may go against the authors' other criticism of Battich et al. (2015). Hansen et al. also mention that Battich's data for nuclear buffering was collected for transient dynamics that was induced by EGF, which was not at steady state, as was considered by the authors. The question then is why would steady-state noise attenuation be better than (or more evolutionarily preferred) than transient noise attenuation? Many genes that need to be accurate (such as in development) are only turned on for a brief period, so noise attenuation in these cases should be only selected for the transient case, not the steady-state case. Thus, the argument here needs to be more carefully constructed.

Miscellaneous Comments

Regarding the statement pertaining to Figures 6 and S6 that, "These multi-state degradation and translation models—where translation and degradation are mutually exclusive—appear necessary and sufficient to explain the amplified mRNA noise in the cytoplasm, as well as the measured protein noise for the various promoters and integration sites examined", we feel that "necessary and sufficient" is a strong logical phrase but "appear" is vague. Is it possible to make this less contradictory?

Summary Remarks

While it is indeed plausible that the conclusions from this paper are correct, it is also plausible that if the authors had used the same mathematical model and extrinsic noise filtering technique as Battich et al. that they would get the same results. At minimum, the authors should show that they recover their noise amplification results with a diffusive model of mRNA export.

Ideally, Hansen et al. should either reproduce the noise attenuation in the cell lines used by Battich et al., supporting the authors argument that the attenuation seen in that paper is an artifact of cell type and condition, or Hansen et al. should reproduce the same extrinsic noise filtering methods used in Battich et al. and show that their noise amplification result is indeed robust to methodology.

At a stylistic level, the authors should go to greater effort to make clear the limitations of their methods, e.g., that there are important modeling assumptions that are made and that there are tradeoffs between different ways of measuring noise in biology. In its current form, the paper comes off as somewhat of an indictment of prior methods, when the reality is somewhat more nuanced and makes one-to-one comparisons inherently challenging. We even suspect that in some sense, everyone is right. Ultimately, a complete resolution of the issues highlighted in this paper seems important and doable—if not immediately, then in follow-up work. We hope this review is helpful, and we are eager to do anything we can to help further. Also, if we have misunderstood anything, we're eager to be corrected, particularly regarding the simulation and theory aspects.

ACKNOWLEDGMENTS

The project was sponsored by the Defense Advanced Research Projects Agency (Agreement HR0011-17-2-0008). The content of the information does not necessarily reflect the position or the policy of the Government, and no official endorsement should be inferred.

REFERENCES

Battich, N., Stoeger, T., and Pelkmans, L. (2015). Control of transcript variability in single mammalian cells. Cell *163*, 1596–1610.

Darzacq, X., Singer, R.H., and Shav-Tal, Y. (2005). Dynamics of transcription and mRNA export. Curr. Opin. Cell Biol. *17*, 332–339.

Fraser, H.B., Hirsh, A.E., Giaever, G., Kumm, J., and Eisen, M.B. (2004). Noise minimization in eukaryotic gene expression. PLoS Biol. *2*, e137.

Hansen, M.M.K., Desai, R.V., Simpson, M.L., and Weinberger, L.S. (2018). Cytoplasmic amplification of transcriptional noise generates substantial cell-to-cell variability. Cell Syst. *7*, this issue, 384–397.