Response to reviews of "Extreme heterogeneity of influenza virus infection in single cells" for *eLife*

Alistair B. Russell, Cole Trapnell, and Jesse D. Bloom

January 3, 2018

Below, the original comments are in blue, and our responses are in black.

SUMMARY

Your manuscript describes the use of a novel single cell RNAseq method to examine heterogeneity in viral mRNA production within single cells infected with single influenza A virions. You report that viral mRNA levels, as a proportion of total cellular mRNA, can vary significantly between infected cells, and this variability cannot be simply explained by the activity of innate anti-viral defenses. Further, correlating the levels of specific host transcripts with viral transcript levels allows you to identify pathways that tend to correlate with viral levels.

The strengths of the paper include the description of a novel method for leveraging single-cell RNAseq to quantify heterogeneity and stochasticity in viral gene expression during infection. This is a really important question for the field, and the approaches detailed here will likely by adopted by other groups studying other viruses. The authors do a good job of accounting for many of the pitfalls inherent in the experimental execution and data analysis. The manuscript is well written, and does an especially effective job of visually presenting complex datasets.

The main shortcoming of the paper is that it provides very little in the way of new information. As the authors clearly and honestly point out, most of the findings in this paper are simply confirming observations made in older papers. Additionally, the findings are purely descriptive, and provide little insight into the mechanisms that may give rise to the observations.

Thank you for the excellent and fair summary of our work. We agree that the approach is powerful, and thank you for recognizing our efforts to fully describe pitfalls and effectively

present the complex datasets. As you note, we agree that our paper does not provide any truly new or unexpected biological result. We do our best to honestly acknowledge this fact in the paper. However, we do believe that the paper provides before the most explicit and direct observations of phenomenon that have previously been mostly inferred from circumstantial evidence. These include the extreme variability in transcripts per cell, the rarity of interferon induction even in the absence of NS, the high rate of missing segments, and the role of co-infection in...

We believe that addressing the points below would help ameliorate some of the shortcomings of the paper.

MAJOR COMMENTS

1. There are some obvious sources of variability that haven't been suitably discussed.

- The first is variation in the timing of the early stages of infection. No steps appear to have been taken to synchronize infection, or to limit secondary spread of the virus within the cultures. Could heterogeneity in the timing of binding/entry/fusion/trafficking explain a lot of the variation observed?

Our argument here would be that washed cells in 8-2 essentially do this. Alistair is generating data by flow with heterogeneity with washing +/- synchronization of infection. The thought is this will show that synchronization doesn't matter with respect to HA protein heterogeneity. This would be a new supporting figure.

The co-infection barcodes provide good evidence against much secondary spread. We can explain that.

- Another potential contributor is the cell cycle status of the infected cells. Did the host transcript data shed light on whether cell cycle status influenced viral transcription levels? This is especially important to address given that some of the genes showing association with the viral burden are cell-cycle related.

Cole had some suggestions of how we could do a few analyses that address this:

This is a super tricky thing to do - I have not seen computational methods that I trust described in the literature for detecting genes that co-vary with cell cycle. Aviv has a method that seems reasonable at first blush but I have some concerns with in terms of the details. I have had better luck testing for genes that are differentially expressed between post-mitotic/quiescent and those that are actively cycling

What would first do is add up the transcript counts for well known markers of actively proliferating cells (CCNB2 is my favorite, but I you could sum a bunch of cyclins and

2

CDKs) to get a ?proliferation index? for each cell. Then look at the distribution of this - I would expect it to be lognormal for a population of cells that are robustly growing. For cells that are mostly quiescent (e.g. a confluent epithelial plate), I would guess the transcript levels for the cells would be low, maybe with a long tail to the right. Then you would add the proliferation index values as a new column in the pData table in Monocle. Then you can use the index as a variable in the model formula for differentialGeneTest. This will give you genes that positively correlate with the proliferation index. The trouble is establishing what values of the index really correspond to actively cycling cells. Ideally, you?d have a bimodal distribution and you?d be able to set some clear threshold, but in practice this doesn?t usually happen. In general, this is pretty tricky to do without built-in controls (e.g. sets of cells that are sorted on DNA content, arrested with some drug, or what have you). So I am generally incline to take these sort of analyses with a huge grain of salt.

- In assessing to what extent lack of RNP expression accounts for the viral mRNA expression variability (page 6, paragraph 2), we think it is important to take into account potential extracellular contamination. Specifically, it would be more convincing if the analysis omitted cells with mixed wildtype/synonymous clones.

If we didn't already do this, we can do it easily — it shouldn't have much of an effect because mixed cells are rare.

We can potentially add a supplemental figure based on wildtype / HA-GFP co-infection.

- As the variability of the viral RNA load appears to be the central result of the manuscript, it would be useful to:

a. Employ a statistic to quantify the variability (e.g. entropy or Gini index)

b. Use simple models to illustrate how much variability can be accounted for by simple effects, such as expected Poisson co-infection frequency, or the likelihood of attaining full complement of RNP genes.

Gini index is something we could definitely calculate. Not clear what we could compare it to, but we can at least calculate it and report it. We then calculate Gini index under some simple model where everything is due to decreased transcription when missing segments. Any such model is probably pretty bogus and oversimplified, so we'd want to do this in a way where everything ends up in a heavily caveated supplement that doesn't bog down text.

2. While the silent tagging method used to address co-infection is clever and appreciated, the issue is not fully settled. The dismissal of co-infection as a factor influencing cell-to-cell variation in lines 127-9 is based on too few cells to draw any conclusions, and thus needs to be tempered. Also, there is likely to be a significant amount of cryptic co-infection with

identical barcode viruses (expected to be similar to that of mixed barcodes, 10%) that could influence measured heterogeneity. These points should be made in the text.

We should make sure that we don't say co-infection is unimportant. We're just saying that low MOI, it doesn't explain everything. Probably this needs to be re-worded.

3. The analysis of host cell transcripts positively or negatively associated with high viral mRNA expression is pretty minimal. Do the host genes identified here match up with the results of previous studies that screened for host pro- and anti-viral factors (reviewed in Watanabe et al. DOI: 10.1016/j.chom.2010.05.008)? Also, targeted gene knockdown or over-expression experiments could help establish the causality underlying these relationships.

In response, gently point out that this isn't actually the point.

It's not that analysis is minimal, it's that there isn't anything.

Then we can somehow work in the Watanabe reference.

We have the oxidative stress thing to add as a figure supplement.

4. All analysis of viral gene expression seems to be at the segment level. How do different transcripts expressed from the same gene segment compare (i.e. NS1 and NEP)?

We have to explain that we can't reliably distinguish NEP/NS1 and M1/M2 for technical reasons that also affect essentially all other single-cell RNAseq.

5. The determination of the minimum required influenza fraction (Figure 4) is based on sound logic; however, we are concerned that the observed results do not fully align with this model. Specifically, while the 10hr experiment looks reasonable, the distributions of wildtype/synonymous mixed cells in other experiments do not appear to show the same trend (most notable for 6hr and 8hr samples, where almost no mixed cells are observed at low fractions). In that regard, using the 10hr dataset to estimate thresholds for all of the other samples does not seem appropriate.

Valid point, maybe we should look at sample-specific cutoffs.

MINOR COMMENTS

- Fig 2: Normalizing high versus low DI stocks based on equal TCID50 to compare IFN induction is not appropriate since these stocks will differ in total particle number. In this experiment, the cells would receive a much higher overall dose of viral RNAs with the high DI stock compared to the others.

Gently explain why this is not relevant for the point we're trying to make.

- Fig. 3B: While I appreciate the density of information presented by this figure, the authors' main point concerning heterogeneity in the viral proportion of total mRNA reads might be made more clearly by simply plotting the relative proportion of mRNAs that are viral per cell.

We can add suggested plot if it isn't already there, it might already be 4E.

- Line 110: Should say "likely reflecting co-infection"

- Lines 186-8: This statement doesn't make much sense since you are talking about variation in mRNA proportion. Global effects such as cell size should not affect proportion.

- Lines 193-8: Is this really remarkable? As the authors point out, there is an extensive body of work demonstrating complementation during IAV infection.

- Line 329: "characterizing" misspelled.

- Lines 330-4: This statement about a confluence of factors leading to extreme cell-to-cell heterogeneity should be reworded since the data here actually doesn't show a clear role for innate immunity in promoting heterogeneity and there is no significant evaluation of the effect of host cell state on heterogeneity.

- The "replicate" 8-hr sample isn't really a replicate since a different MOI was used. Also, the purpose of the differences in treatment for this sample compared with the 8-hr sample are not well explained.

Good point that we'll address already in first comment.

-Figure 3 is not effective in supporting the observations made in the text (that most cells have little viral RNA, and that there's a wide variability in the amount of viral RNA). Separate distributions for the viral RNA would be more telling. Log scale would be more appropriate for showing counts per cell.

-Figure 4C: the threshold line appears to be a bit to the right of the global maximum - going through the description, it was not clear why that should be the case.

-Why was media change skipped in the replicate experiment?

-Consider using different color schemes for Fig 4A and B, as different variables are being shown.

-Figure 5A legend: axis descriptions appear to be swapped