**EDITOR COMMENTS:**

Thank you very much for submitting your manuscript "Single-cell virus sequencing of influenza infections that trigger innate immunity" (PPATHOGENS-D-18-02247) for review by PLOS Pathogens. As with all papers submitted to the journal, yours was evaluated by an Associate Editor in consultation with the Editorial Board. Your article was also evaluated by independent reviewers.  All three reviewers appreciated the use of cutting edge technologies to allow viral genome sequence and the host transcriptome to be analyzed in parallel at a single cell level. However, the reviewers also consistently noted that novel insights into influenza virus biology and virus-host interactions were lacking. Based on the reviews, I regret that we will not be able to accept this manuscript for publication in the journal.

Thank you for summarizing the rationale for the *PLoS Pathogens* decision. While we are disappointed that the reviewers did not find the paper to have novel biological insights of sufficient impact for *PLoS Pathogens*, we appreciate that they found it methodologically interesting and scientifically sound. We have chosen to transfer the manuscript and reviews to *PLoS One* using the option described here: <https://blogs.plos.org/plos/2017/10/manuscripts-take-flight/>. Based on the *PLoS Pathogens* reviews, we think the manuscript is suitable for *PLoS One*, which explicitly states that it accepts “scientifically rigorous research, regardless of novelty.”

The reviewers did have several major and minor comments, all of which we have addressed. Below we provide a point-by-point description of our revisions in response to these comments. We have also provided a revised manuscript with tracked changes.

**REVIEW COMMENTS**  
  
**Part I - Summary**  
  
Reviewer #1: This manuscript by Russell et al uses a very powerful single cell approach to determine both the mutations in the viral genome by pacbio and the antiviral response by 10X. To hone in on cells that have sensed virus replication the authors construct a MACS sortable ISG reporter cell line. The authors demonstrate significant heterogeneity in both the induction of IFN and the viral genome. Overall the manuscript is well written, the figures well constructed and experiments well controlled. While the major finding of the manuscript, that lack of NS1 associates with increased IFN production, has been widely demonstrated the powerful technique and careful analyses performed here will be of significant interest to the field.

We thank the reviewer for praising the paper as “well written”, the figures as “well constructed” and the experiments as “well controlled” experiments. We hope that (s)he is correct that the “powerful technique” will be of “significant interest to the field.”  
  
Reviewer #2: This manuscript by Russell et al builds upon their previous publication in eLife to examine the relationship between viral genetic diversity and the seemingly stochastic patterns of IFN induction during influenza infection. To do this, they develop a novel viral sequencing approach based on PacBio long reads of viral mRNAs isolated from single cells. They also develop a clever reporter cell line that allows them to enrich cells positive for IFN beta or lambda promoter activity. Using these new tools, the authors define and compare the complete viral genome sequence and host transcriptomes from 150 infected cells. They find that roughly a third of cells are infected by a completely “wild type” virus, and that the remaining two thirds have some sort of SNP or deletion. This approach yields some interesting findings. The graph of the individual viral genotypes across all 150 cells is a nice representation of viral genetic diversity. As to the central question, they note an increase in the frequency of IFN induction in cell infected with mutant virus versus those infected by WT (30% vs. 20%), but this difference does not quite achieve statistical significance. They do find significant associations between IFN induction and specific mutations in the NS and PB1 segments. They go on to generate recombinant versions of these mutants and confirm that they trigger IFNbeta reporter induction more efficiently than wild type virus.  
  
Overall, the paper is very well written and the authors do an excellent job of laying out detailed methods for the new single cell viral sequencing approach they develop. These new methods/tools may be of use to other groups in the field. The conclusions are generally well supported by the data, though there are a few instances of over-interpretation. The paper does a good job of clearly illustrating that wild type virus stocks harbor genetic variants that can encode non-wild-type phenotypes, including enhanced IFN stimulation. This is an important point, and though it largely follows from the work of many groups describing the genetic variation present in influenza virus populations, it has never been examined in this way before.   
  
The main weakness is the lack of follow-up beyond the initial identification of a handful of mutations associated with enhanced IFN induction. The primary hit is a loss of NS1, which could be seen simply as a positive control given the vast body of work detailing the role of NS1 in antagonizing IFN. The second hit is a DVG, thus again confirming what other groups have reported. The remaining few hits are point mutants in PB1 and NS1 that are not explored further and only have a modest (2-3fold) increase in IFN+% compared with wild type. Based on the limited number of hits and absence of follow-up, we learn nothing new about mechanisms of IFN induction during influenza infection, the central subject of the paper. Thus, this paper is really more of a methods paper than the report of a new scientific discovery.

We appreciate the summary of our work. We are especially gratified that the reviewer considers the conclusions to be “generally well supported by the data,” and that the “new methods/tools may be of use to other groups in the field.” We recognize the reviewer’s point that our work is more a methods paper since the biological findings are not entirely novel. This concern should be alleviated by the fact that we are transferring the manuscript to *PLoS One*, which simply requires the work to be scientifically sound. As detailed below, in the revised manuscript we have addressed the few instances where the reviewer found the results to be over-interpreted.  
  
Reviewer #3: Innate host responses to viral infections (here influenza A virus)  in a given pool of host cells vary in amplitude and quality from cell to cell. The authors aim to unravel this heterogeneity by combining transcriptomics approaches with PacBio sequencing of viral mRNAs from single host cells. Their results allow them on single host cell level to correlate alterations in the viral genome (expression levels and mutations in viral mRNAs) with the quality of interferon and ISG responses.

Per se this is an interesting and novel approach to identify viral factors determining the antiviral host response, which might not be detectable in bulk approaches. However the manuscript is largely based on evaluation of RNA based readouts, yet the arms race between host and virus clearly takes place on protein level. It remains elusive, how the authors want to control for presence or absence of a given viral or host protein underlying their claims. Further the manuscript leaves the reader puzzled about the take-home message, since the only significant finding presented is that NS1 deficiency significantly correlates with IFN induction, which is fairly well established.

We thank the review for noting that we have described an “interesting and novel approach.” We agree that NS1’s role in IFN induction has been described before, and are transferring the manuscript to *PLoS One* where the emphasis is on scientific soundness rather than perceived impact. We have addressed the reviewer’s other specific comments as described below.  
  
--------------------  
  
**Part II – Major Issues: Key Experiments Required for Acceptance**

Reviewer #1: My major concern is that some of the results are over interpreted. The authors state several times that defects in viral genomes increase the likelihood of the cell producing interferon. However, the authors show no statistical data to support this claim (figure 5b). The authors need to amend the statements in the abstract and summary to more accurately reflect the data.

This is a completely valid point. Most cells that produce IFN in our experiments are infected by mutant virions, but the numbers are too small to make the trend of mutants producing more IFN statistically significant—although the validation experiments do show that some of the mutations are IFN-inducing.

We have ensured that there is text in the Results that explicitly states this fact: “Cells infected by incomplete or mutated virions also expressed IFN more frequently than cells infected by wild-type virions (Fig 5B), although this difference was not statistically significant (P = 0.12, Fisher’s exact test)… This lack of statistical significance after FDR correction could be due to the relatively modest number of fully sequenced infected cells (just 150). The validation experiments in the next section show that many of the viral mutations in IFN+ cells do in fact increase the rate of IFN induction.”

As the reviewer suggests, we have re-worded both the abstract and summary to ensure that they simply state the two facts from our data: (1) most cells that induce IFN are infected by mutant virions; (2) some of the viral mutations found in IFN+ cells increase the rate of IFN production (validated by the experiments in the last section of the results). We have ensured that the revised abstract and summary do not claim that there is a statistically significant association between the presence of viral mutations and the rate of IFN induction.   
  
Reviewer #2:

-  For this paper to be something beyond a methods paper, it would need to leverage the really novel dataset they produce to generate, and ideally test, a new hypothesis. For instance, is anything known about the specific residues in PB1 that could suggest why they result in increased IFN induction? Testing the effects of these substitutions on polymerase function could help shed light on this.

We have chosen instead to maintain the more methodological focus and transfer the manuscript to *PLoS One*. As the reviewer notes, this makes the paper more of a “methods paper.” We have made revisions to clearly emphasize this fact: for instance, in the revised abstract we now say: “Here we develop a new method to…”   
  
-  While many of the specific analyses are highly quantitative, the overall conclusions about the contributions of viral diversity to IFN induction are not. While they do show that mutations can be associated with elevated IFN induction, it is unclear whether mutation-dependent induction is a significant contributor to IFN induction by the population as a whole. A quantitative assessment of the relative contribution of the deterministic effects of mutations to IFN induction (versus the inherent stochastic behavior of the induction pathway) would strengthen the significance of the paper. This could possibly be done by quantifying the effects of increasing the random mutation load on IFN induction using a mutagenic drug.

This is an accurate point. We agree that new experiments such as the one reviewer suggests would be interesting, but they are beyond the scope of our current study. Instead, we have made the revisions described in the response to Reviewer 1 above. As described in that response, these revisions ensure that our manuscript clearly states what we have and have not demonstrated about the role of viral mutations in IFN induction.  
  
Reviewer #3:

1) The authors find a number of viral mRNA variants correlating with a higher IFN mRNA level. This includes deletions of whole segments and or in/dels in single genes as well as non-synonymous substitutions. The authors show that only the NS1 deletions seem to severely impact the correlation between viral RNA and IFN mRNA expression (in presence of NS1 no correlation, in absence IFN levels correlate positively with viral mRNA levels). Why are the levels of viral mRNA in IFN negative cells lower, when infected with an NS deficient virus (5D)? If NS1 were the main IFN antagonist, one would expect that a virus lacking NS1 replicates well in IFN negative cells (unless they are ISG positive). And why is NS1 not doing anything in the IFN+ cells in the upper panel? 

The reviewer is commenting on the fact that Fig 5D shows that in the single-cell data, only for viruses that lack NS1 is higher viral gene expression associated with a cell being IFN+.

First, we have added new data (Fig S14) that confirms this basic finding using flow cytometry. These new data show that in cells infected by a virus with mutations in NS1 (but not cells infected with other viral variants) that higher gene expression (measured by HA levels in the case of the flow cytometry) are more likely to be IFN+. These new data corroborate the based result in Fig 5D by an independent assay. These data are discussed in the last section of the results of the revised manuscript.

We have not worked out the mechanism. But we speculate that the effect has to do with the probability that an infection activates IFN, not due to the anti-viral activity of ISGs. In our experiments, all infections happen prior to IFN induction since we allow just a single cycle of infection. We speculate that NS1-defective virions are unable to block IFN induction by new viral products, and hence there is an association between viral gene expression and IFN in these cells. We speculate that for NS1-expressing virions, increased gene expression also makes more NS1 that antagonizes IFN production. However, this remains speculation so we have not added it to the manuscript. But certainly the differences between NS1+ and NS1- infections indicates the power of examining IFN induction on the single-cell level.

The reviewer also asks why NS1 is “not doing anything in the IFN+ cells in the upper panel.” One of the results of our work is to show that sometimes IFN can be induced even in cells that express NS1; for instance, Fig 4 shows a number of examples of such infected cells. Therefore, NS1 is clearly not *perfect* at antagonizing IFN.

2) The data presented in figure 4 are puzzling in a number of ways:  
  
        a) viral mRNA levels do not seem to depend on the presence or the integrity of viral polymerase genes (cell 2, 3, 5, 17 and others). What explains the high content of viral mRNA in these cells (up to 60% of total  RNA)?   
        b) Why are the two virus variants so unevenly represented?  
        c) How much of the genetic heterogeneity is a consequence of the IFN response on the viral replication machinery? Would IFN receptor deficient cells show the same distribution?  
        d) Omission of transcription of one of the segments occurs quite frequently with exception to NP. Could the authors speculate why this is?

These questions are due to the reviewer mis-understanding Fig 4, which was our fault as the original legend was ambiguous. The light and dark blue colors in Fig 4 reflect whether the cell was infected by one viral variant or co-infected by both viral variants---*not* which viral variant infected the cell. We have revised the legend to the figure and the figure labels to make this fact clearer. We have also added some additional text to the results that makes this point clearly.

Once the meaning of the colors in Fig 4 are understood, the answers to the reviewer’s questions become obvious:

1. All the cells that the reviewer refers to as lacking in intact polymerase gene are *co-infected* with two viral variants (hence the dark blue bars). In all these cells, one of the viral variants has a deletion in the polymerase, but the other viral variant has a full-length copy (hence the fact that the yellow lines are only half the height of the bar). Therefore, these cells are still expressing full-length versions of all of the polymerase genes.
2. The predominance of light blue over dark blue bars simply shows that most cells are *not* co-infected by both variants, as is expected due to the low MOI that we used. The relative frequency of the two variants is shown in Fig 3.
3. We think that very little of the viral genetic heterogeneity is due to the IFN response since (as shown in Fig 5D and discussed in response to reviewer’s point immediately above) there is no tendency of the IFN response to reduce viral transcription, at least at the early timepoints in our experiments. Repeating all the experiments in IFN receptor-deficient cells would be interesting, but is beyond the scope of our current study.
4. We offer two explanations for this fact in the legend to Fig S5, where we show complete data on the rate of segment absence in our experiments. Specifically, we say: “The exception is NP, which is detected in virtually all infected cells. The much higher frequency of detecting NP could reflect a biological phenomenon, but we suspect it is more likely that cells lacking NP tend to have much lower viral gene expression overall and so are not reliably called as being infected in our experiments because the number of viral mRNAs is below the detection limit.”

3) Would the mutant viruses tested in bulk cells (Fig.6) show the same heterogeneity on single cell level?

Yes, they do show the same heterogeneity at the single-cell level.

We have now added a new paragraph to the last section of the results that discusses this fact. It says in part: “Fig 6 shows flow cytometry data (see also Fig S12), which is itself a single-cell measurement, albeit one that does not report the viral genotype. As can be seen from these data, none of the mutant viral stocks induce FN in more than 20% of infected cells… Therefore, the experiments in this section not only validate some specific viral defects that increase IFN induction, but also show that induction remains stochastic even with those defects.”

--------------------

**Part III – Minor Issues: Editorial and Data Presentation Modifications**

Reviewer #1: The key for figure 4 should include the green and orange boxes to make interpretation easier.

The revised Fig 4 now includes a key for the green and orange scales found in the boxes.

The authors should include titer data for the viruses described in figure 6. This may help with interpretation. E.g. Were the mutants that failed to induce IFN also fail to grow to high titer?

Alistair, do you have some data that you can mention here? I’m guessing it’s at least somewhat confounded as (for instance) the PB1del had to be grown in complementing cells. But at least for all the point mutants and perhaps for NS1stop we can say what the titers were post-rescue or post passage? If we don’t want to emphasize this fact since we didn’t do exhaustive titering, we could even just list these numbers in Methods and refer reviewer to that.  
  
Reviewer #2:

-  It wasn’t really clear how many viral genomic RNAs were actually being successfully sequenced in each cell, and how the decision to call a gene segment as missing versus just undersampled was made. Maybe this was buried somewhere but I didn’t see it. A clear, concise description of the rules used to makes these calls as well as data on the distribution of individual viral cDNA molecules per cell that made it into the analysis would be useful.

The method used to call gene segment presence / absence is summarized in the third section of the results in the paragraph beginning: “We called the presence or absence of each viral gene in each infected cell, again using a Poisson model parameterized by background fractions estimated from uninfected canine cells…” As explained in this paragraph, we used the un-infected cells to estimate how much signal we would see just due to background (there is always a bit of background in single-cell RNA-seq), and then examined if there were more mRNAs for the gene than expected from this background. The full details are in the Methods.

The cell-gene matrix included at the GEO link in the paper gives the exact counts of each viral and host transcript in each cell. These data are summarized in Fig 3B (total viral mRNAs detected per cell) and Fig 3E (total fraction of viral mRNA from each gene for each cell). We are not able to create a more compact representation than those figures as plotting the raw counts for each cell separately is visually uninterpretable since there are so many points spanning such a wide range.  
  
-  Lines 237-9: the observation about higher viral gene transcription levels associated with IFN induction in absence of NS1 is interesting but is just kind tossed off with no further discussion. Any thoughts on what could be going on here?

See our response to the first major comment of Reviewer 3. We do not have a complete mechanism, but have added new data that use an independent method to show that this trend is real.  
  
Reviewer #3:

1) The authors start from 1614 cells and filter out 124 cells due to high or low transcript levels, likely indicating multiples. It remains unclear, why the low transcript samples were filtered out in this context.

Filtering low-transcript cells is a standard recommended step in single-cell RNA-seq analyses, since it gets rid of low-quality emulsions (for instance, see Haque et al, Genome Medicine, 9:75).

We now explain this fact. Specifically, we have added the following sentence: “To remove some of these multiplets along with low-quality droplets, we filtered transcriptomes with unusually high or low numbers of cellular transcripts as is commonly done in the analysis of single-cell RNA-seq data [Haque et al, Genome Medicine, 9:75].”