Here are the *eLife* reviews. I don’t think these are particularly fair or helpful. The reviewers apparently don’t find it interesting, and I simply disagree with their reasoning so nothing to do there.

They seem to complain that a lot of the data is not “new”, but this is simply incorrect. In Figure 3, we make some plots that are similar to the earlier paper—but it is a new dataset, and it’s important to re-plot these quantities for the new data set. So I don’t think we should get rid of that, as it would seem like scientific negligence to just assume without validation that everything that was found in the original paper is also going to be true for the new dataset here.

They also make a number of specific critiques. I’ve put point-to-point responses to these below (just to guide our thinking on what to change). I feel that in the majority of cases they are asking about things that we already clearly state, so I’m not sure how much to change. But in some cases I guess we could state a bit more clearly.

One of the reviewers also makes a big deal about the MACS, for instance wanting experiments that show that the MACS enrichment is linear and reproducible. But I don’t think this is worth doing. None of our conclusions depend on the MACS enrichment being linear (I’m not even sure what they mean “linear” with respect to). We already have a supplemental figure that shows how the MACS enrichment works in two separate replicates, and I don’t think it is expected that the exact enrichment percentage of MACS will be identical over different days. Anyway, we just use MACS to get more IFN+ cells, and all the downstream conclusions are based on analyzing these cells with built-in internal controls (such as the canine and uninfected cells).

I am in agreement with your conclusions. I am very confused about the focus on MACS as well, given that it is more-or-less confirmed via several orthogonal means (qPCR, response to Sendai, response to NS1mutant virus, consistency with prior reporters developed by Randall’s lab, and the success of our single-cell experiment itself! I don’t think I have ever seen a method as confirmed as we did!) I do think emphasizing our new data could aid in future reviews—for instance others have posited that remaining stochasticity in an NS1 mutant is due to defectives, or posited ALL interferon is driven by defectives, whereas here we very clearly rule out a wide swath of prior speculation, which I think has significant value (ie we do not comprehensively describe all immunostimulatory heterogeneity, but we do conclusively rule out facile explanations attributing all variation to a single genetic (or non-genetic) source).

Dear Dr Bloom,  
  
Thank you for submitting your work entitled "Single-cell virus sequencing of influenza infections that trigger innate immunity" for consideration by eLife. Your article has been reviewed by 2 peer reviewers, and the evaluation has been overseen by a Reviewing Editor and a Senior Editor. The reviewers have opted to remain anonymous.  
  
Our decision has been reached after consultation between the reviewers. Based on these discussions and the individual reviews below, we regret to inform you that your work will not be considered further for publication in eLife.  
  
We are sorry to say that, after consultation between a Senior Editor and two other experts, we believe that your paper does not reach the level of significance that we seek for a Research Advance. While we continue to believe that your technology is compelling and the findings emerging from it are important, almost all of the interesting new data in this paper is only in figures 5 and 6. Thus, we feel that the current manuscript is better suited for publication in a specialty journal.

This doesn’t seem true with respect to new data. Figure 1 is new data, Figure 2 is a schematic of a new approach, and the data for Figure 3 shows some similar plots (and also some new ones) to our prior paper but for a new dataset. Figures 4, 5, and 6 are all new.

Agree. Feel we can avoid these misconceptions as stated above.  
  
eLife is highly selective, which means that the majority of submissions are rejected, but we thank you for sending your work for review and we hope you will submit to eLife again in the future.  
  
Best wishes,  
  
Arup Chakraborty  
Reviewing Editor  
  
Arup Chakraborty  
Senior Editor  
  
Reviewer #1 (General assessment and major comments (Required)):  
  
This manuscript builds on a previous publication by these authors that harnessed single cell RNAseq to describe the extent of heterogeneity in viral gene transcription patterns between individual infected cells. The authors extend the methods employed in the previous paper by performing long-read sequencing to reconstruct full length viral mRNA sequences within single cells. Additionally, they develop a clever technique to enrich for cells based on IFN promoter activity, a critical advance given the rarity of flu infected cells that produce IFN.  
  
The overall goal of the study is to define the contribution of viral genetic diversity to heterogeneity in the cellular IFN response to influenza infection. The authors reconstruct the complete viral genomes from 150 infected cells and attempt to correlate the occurrence of different viral genetic lesions (SNPs, deletions, and gene loss) with the presence or absence of an IFN response. They observe that cells infected by virions that either lack the NS gene, or carry SNPs in the PB1 gene are significantly more likely to be IFN+. They then use recombinant viruses to confirm that a handful of these variants significantly increase the frequency of IFN induction.   
  
Overall, the paper is well written and many of the conclusions are supported by the data. The approaches taken here represent a real technical advance, and the authors do an excellent job of thoroughly detailing and addressing many of the technical issues involved and providing excellent documentation for others who may be interested in performing similar analyses. The broader question asked here is an important one, and the authors are to be commended for their ambition in tackling this challenging and important issue and pursuing it with much greater sophistication than has been done previously.   
  
My enthusiasm for this manuscript is significantly dampened by the fact that the results do not provide any new information or insight. The primary take-home message, that viral genetic diversity (primarily in the NS1 gene) may contribute to heterogeneity in the IFN response is highly predictable, especially when compared with the alternative hypothesis: that viral genetic diversity has no effect on the host response. Further, the authors do not actually demonstrate that viral mutants are a major driver of the IFN response, as the difference in IFN induction between WT and mutant virus infected cells was small. The primary genetic lesion identified as a driver of the IFN response, loss of the NS gene, simply confirms the multitude of studies over the years that have demonstrated that NS1 is a potent antagonist of the IFN system. The relatively tiny number of SNPs outside of NS1 identified as inducing IFN are too limited in number to provide any broader mechanistic insight.

I don’t really agree with this assessment. There are no prior papers that directly show that viral genetic diversity is associated with IFN induction. Although many papers have posited that viral defects could enhance IFN induction, there are no others that actually look to see which defects are found in IFN+ cells. We do directly demonstrate that the viral mutants we identify increase IFN response, validating a number of them in Figure 6. The reviewer complains that the increase in IFN induction by these mutants is “small” (in some cases, it is 10-fold!), but this is also a finding—it shows that no mutations deterministically induce IFN.

Yeah, agree strongly. 2-fold differences in point mutants have been the topic of entire manuscripts (including a recent Science paper), and we describe phenotypes that are more severe here. Moreover, we not only identify absence of the NS **segment** (not just point mutations) as a feature driving IFN induction, but found it is **not necessary**. A nuanced view of IFN induction has been sorely lacking form the field, and naivete regarding a single, uniform, path to induction has prevailed.   
  
Part of the difficulty here lies in the relatively small number of cells examined, coupled with the high degree of stochasticity inherent in the system. Technical limitations (not the fault of the authors) restrict the number of cells that can be examined, and as a result the study too often seems to be offering anecdotes about handfuls of infected cells, rather than providing a compelling quantitative dissection of the issue. Overall the manuscript does not significantly advance our understanding of the interplay between viral diversity and the IFN response.

I agree it would be nice to have more then 150 cells, but we’ve gone from knowing the sequence of the infecting virions in zero IFN+ cells (this has never been measured before) to 150. Seems like a valuable start. And it is not just anecdotes, we validate a number of the mutations in Figure 6.

Agree. I still think the most significant point is we rule out a number of hypotheses regarding interferon induction by influenza. While anecdotal with respect to total immunostimulatory breadth, our data show that simplistic explanations of interferon induction are simply incorrect. Explanations that, mind you, have still been posited in recent publications, and therefore our work represents a significant correction to current models of innate immune stimulation by influenza.  
  
Specific points:  
Figure 6: It would be nice to back up the flow cytometry data in figure 6 with RT-qPCR analysis of IFN beta/lambda induction to see how well the single cell reporter data here matches up with bulk transcript numbers.

We could do this, although we extensively validate the reporter. The main reason to not do this is that qPCR is sensitive to infecting viral dose, so we’d have to get that exactly normalized.

Agree. Other reports are plagued by particle miscounts, which are largely corrected by this approach.   
  
The authors show that a single DVG identified in their sequencing can induce IFN and point towards this as confirmation that of the widely held idea that DVGs induce IFN; however, many other cells with viral deletions do not seem to induce IFN induction. Can the authors weigh in on this discrepancy?

This is one of the main findings of our paper: no mutations deterministically induce IFN. We say this multiple times, for instance: “Additionally, none of the viral defects in IFN+ cells induce IFN deterministically in our validation experiments—and the apparently most immunostimulatory defect (absence of NS) also occurs in multiple IFN- cells in  our single-cell dataset.”

We could put back in a nonstimulatory particle if we wanted to.

Reviewer #1 (Minor Comments):  
  
Figure 1C: I understand that the two were often highly correlated, but please specify here whether these data show are type I or type III IFN+, rather than just IFN+

The figure supplement says that this is type III IFN, but we can add to the legend of Figure 1C.  
  
Figure 2: sections B-D are not very clear. As I understand it, each color represents a cell-specific barcode. Why then are all polyA transcripts from the red cell (which presumably contains both viral and host transcripts) subject to full length sequencing, but none of them form the purple cell?

This is because we enrich for flu genes for the PacBio sequencing, and in our schematic only the red cell is influenza-infected. We can add an annotation indicating this.  
  
Fig 3 supp 2: this would be better presented on a log scale

It is on a linear scale because it’s not possible to show values with zero on a log scale. We could make a log version too and add a pseudocount, but it’s not obvious to me why a log scale would be preferred over a linear one here. The correlation between type I and type III IFN is robustly supported by multiple lines of evidence.  
  
Figure 6: Based on the supplemental data, it seems that these plots show the % of infected cells that are IFN+, rather than the % of all cells. This point should be made clear in the figure legend.

This is true for Figure 6A, and we can state this directly in the legend.  
  
Lines 201-210: The rules for calling a variant were clear, but the rules for calling presence or absence of a viral gene were not. These may be buried somewhere in the methods, but should be made clear without the need to dig.

How gene absence was called is clearly stated in the results. The paragraph beginning on line 146 says: “We then 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.”  
  
The section in the methods detailing the methods for Pacbio library prep for flu clearly reveals that this was a heroic effort, but could be compressed to focus on what worked and briefly detail what didn't to guide other groups interested, rather than providing a narrative of all of the work required to get the final product.   
  
This is a new approach, so we feel that it is important to provide comprehensive methods. In addition, the final dataset uses the sequences from all of the library prep methods, so just describing one approach would not adequately describe how the data were generated.  
  
Reviewer #2 (General assessment and major comments (Required)):  
  
This paper and the accompanying paper, both by Russell et. al, describes a single cell RNA-seq analysis of mammalian host cells in culture infected with Influenza. While the accompanying paper describes heterogeneity of viral transcripts within infected cells, the main advance of the current paper is a cutting-edge technology to analyze at single cell resolution both the induced host response and full-length viral transcripts. The paper represents a very interesting concept and technological advance: the authors develop an enrichment protocol for cells that induce the Type-I/III IFN response. Next, within infected cells, the authors analyze by PacBio the viral transcripts to obtain full-length isoforms. Through this analysis the authors link the presence of mutations in certain viral genes to the induction of Type-I/III IFN response of same infected host cell. While the findings and technology are very compelling, I have concerns that the current paper cannot stand on its own.  
  
Major comments:  
1. My major concern is that the authors are presenting very similar findings in the two different papers. While it seems that the authors did use data from different experiments, they do replicate the figures between the papers. Figure 3D and 3E are exactly the same figures as in the accompanying paper. Figure 3B and 3G are not the same as in the accompanying paper, but they are redundant and used to make the same conclusions. Figure 2E is also very similar. Since figures 1,2 are mostly illustrations, figure 4 could be a supplementary table, almost all of the real interesting new data of this paper remains only in figures 5 and 6. This overlap between the two papers seems disproportionate.

There is no overlap in the datasets between the two papers. We agree that Figures 3D and 3E have the same format as the original paper, but they show a different dataset. Since this is only the second study to examine these issues, we think it is important to show what the data look like for the new dataset. Similarly with Figure 3B—this is a new dataset, so we need to show what the data look like. Figure 3G is not redundant with any other figure in this paper or the earlier paper, as there are no other plots that show the distribution of IFN across cells.

Overall, it is true that we re-analyze some of the same trends as in the earlier paper. However, this is important! What if things were not the same between experiments, we’d want to know that!

2. The authors present a reporter for the Type-I/III activity. However, there are no experiments that show a validation of this system. The authors only show FACS plots of the infected cells (Sendai and Influenza, Figure 1 and supplements), but this does not confirm the sensitivity or the accuracy of their system. This is most important for their follow-up experiments, where the authors rely on their enrichment method that seems to be very noisy.

a. The authors state in the current paper that IFN+ represent 0.5% of infected cells. However, in the accompanying paper, they show only 1 cell with detectable interferon in their data, which represent much lower frequency - can the authors explain this discrepancy?

In the Figure 1C, we report 0.5% of cells are IFN+. In the original paper, we found 1 / 368 = 0.3% cells IFN+. Given experiment-to-experiment variation and the fact that we are using entirely different methods to call IFN+ (flow in Figure 1C and single-cell transcriptomics in the prior study), I’d consider these fairly similar numbers. But we can add a sentence emphasizing that they aren’t completely identical.

I don’t think we really need to change this…..it would be inappropriate for anyone to draw statistical conclusions from one observed event.

b. In line 179: the authors claim that of the uninfected cells, 1.3% were IFN+. This means that 16 cells had IFN+ signal in their uninfected cells. This number is very high for a background. This is especially worrisome when compared to the 20% of the infected, which represents only 58 cells. I am not sure why the authors chose to represent these as percentages; the authors should clearly state the number of cells in each case. In any case, looking at the numbers it seems like a very noisy enrichment process. Several validation steps needs to be introduced into their analysis. For example - the authors demonstrate the efficiency of the enrichment process (Figure 2 - figure supplement 1), but they should compare the enrichment process in infected and uninfected cells. They should also demonstrate the enrichment using a clean setup - spike varying amounts of infected cells into uninfected and show how linear the enrichment is.

The numbers are clearly stated in Figure 3G: 60 of 290, and 15 of 1200. However, we can re-state them in the text too when we refer to the percentages in the figure. We clearly explain why there are some IFN+ but uninfected cells: “Few uninfected cells were IFN+; the few that were present might be because the MACS enriched for rare cells that spontaneously activated IFN, or because some cells that we classified as uninfected were actually infected at low levels.”

We provide a detailed figure (Figure 2 – figure supplement 1) showing how the enrichment process works. This figure shows the final IFN+ and IFN- cells after enrichment, which seems to be exactly what the reviewer is asking for. I don’t see how the other suggested experiments are relevant. None of our conclusions rely on quantitative enrichment: all the conclusion are based on the single-cell RNA-seq after enriching these populations, so why does it matter if the enrichment is linear? In any case, it is well known that MACS enrichment is not perfect, but simply causes a fold enrichment.

Agree. It is unlikely the 16 cells are ‘background’, and are bone fide interferon-producing cells. We would expect no ability to observe enrichment in uninfected cells—there would be no control for gating in flow cytometry, and transcript levels are already below the limit of detection for interferons in those cells. Additionally, as above, linearity of enrichment is unnecessary. I would add that the validation we perform (unimodal peak, qPCR, response to Sendai, response to influenza) is completely consistent with accepted validation for other, similar, reporters that have been used to great success in this field. I would add, if the enrichment was nonlinear…it would not impact the utility of the approach or our results, as you say above.

3. The authors use viral transcript numbers as a direct measurement of viral burden (e.g. line 216). However this concept needs to be experimentally validated directly. This is most obvious in their Figure 6, where the authors validate the effects of viral mutations on IFN+ cell numbers. The authors should have also checked the effect on viral burden in these conditions.

I don’t understand this comment. The reviewer is correct that for the RNA-seq, we use viral transcript counts as a measure of viral transcriptional burden (e.g., line 216). I’m not sure what the reviewer wants us to validate about this? RNA-seq directly measures transcript abundance, and it is well established to use those measurements. How would we validate them?

In Figure 6, the measurements are made by flow cytometry. It is not possible to measure viral transcripts by flow cytometry. Flow cytometry measures protein expression, not transcript numbers.

I would add we include full flow cytometry profiles for Figure 6, including viral protein staining.

4. In line 234 and Figure 5 - the authors claim that viral defects are major contributors to heterogeneity based on number of viral mRNAs. The distribution of the mutated viral particles looks more bimodal then heterogeneous. If that is the case, another option the authors should check is that some of the mutations increase these numbers and some don't. the authors should check what are the common mutation in the higher part of the distribution and in the lower part.  
  
This is what we do. Figure 5B shows that some cells with mutated virions are high in IFN. Then Figure 5C looks to see which mutations are present in those cells. Then Figure 6 validates those mutations actually induce IFN. A key point, however, is that there is **not** a single common mutation that explains IFN induction, but rather many different mutations.

Reviewer #2 (Minor Comments):  
  
Minor comments:  
5. In line 162 the authors claim: "despite the wide variation in absolute expression of viral genes, their relative expression was fairly consistent (Figure 3E)". This is a similar statement to the one they make in the accompanying paper. Clearly, the authors should show that this is not simply due to sequencing coverage (in that case, the same trend should be evident also for host transcripts). Alternatively, if the underlying reason for this inconsistency between relative and absolute expression, the authors should normalize each cell by UMI count (this is not mentioned in their methods section).

It is true that we repeat the finding from the original paper that relative expression of viral genes is similar. We think it is important to show that this earlier finding holds on the new data set.

I don’t understand the suggestion about sequencing coverage. We do use UMIs for all quantifications. I don’t understand what trend we would look for in host transcripts, or what we would normalize by UMI count. Right now the figure shows the fraction of all viral transcripts (counted by UMIs) that come from each viral gene.

I also do not understand what trend they are thinking in the host counts. I think the misunderstanding here is in that they believe we are summating flu reads rather than counting fraction of total reads from a cell? I admit this makes no sense with any interpretation I can surmise.

6. In Figure 3 - figure supplement 4 the authors show a t-sne plot that demonstrates that IFN, and ISG genes contributes substantially to the structure of the data. This is different from the t-sne plot in the accompanying paper, where no such structural feature could be found. The authors should explain this discrepancy.

They are different datasets. In the earlier dataset, just one cell was IFN+ so IFN and ISG did not explain much of the data. We could make this clear in the legend to that figure?

I think this further demonstrates a misunderstanding about which data were new unfortunately. Different data, different populations, different structure…

7. In line 106, the authors mention that they added non-enriched cells to ensure the presence of IFN- cells. Since after enrichment only 20% of the cells are IFN+, I am not sure why this is needed? If the authors would have done optimization and validation of the enrichment process they would have already known that.

We did do optimization and validation of the enrichment process (Figure 2 – figure supplement 1), but as that figure shows there was always some variation day-to-day (this is common for enrichment experiments). Therefore, we added back non-enriched cells just in case the enrichment worked exceptionally well on the day we did the experiment. We can mention this, but in any case it doesn’t seem like adding a few uninfected cells causes any problem.

We actually added infected cells back in, which makes sense. Our enrichment showed that we got enrichment, and that it was not 100 percent of the cells **reporter** positive. We did not stain for influenza, meaning we did not know how many IFN negative flu positive cells there were (which would vary from experiment to experiment contingent on population composition and staining the low numbers of cells obtained from this sorting would have been technically challenging), and, moreover, flow cytometry is defined by decisions made in terms of voltage and gating and it is possible we would have miss-called the percentage and in actuality more cells could be IFN **transcript** positive. Optimization and validation, as was performed, cannot truly rule out those technical issues, but simple remixing of the sample is a robust solution.

8. The process of doublet omission is unclear. In line 130 the authors indicate that they obtained 12 transcriptomes that are a mix of human and canine. Using this number they estimate 11% doublets - how do these numbers fit? As they had spiked only 50 canine cells it seems that the number should be much higher. The authors then exclude doublets only based on UMI count. How was the canine spike useful then? This analysis should be better explained or omitted.

The calculation used for the multiplets is in Bloom (2018). The canine spike-in allowed us to estimate the doublet rate and also estimate the background rate of flu.

9. Line 124 - how do the authors validate their claim that most cells are infected by 1 or 2 virions?

This is done in Figure 3F. We can reference that.

10. In line 139: "as expected only a small fraction (0.7%) of transcripts in the uninfected canine cells were from influenza...". Why is this expected? How do these transcripts distribute between the cells? Is it a small fraction from many cells or large fraction of some cells?

Because the cells are not infected, we expect only the background ambient amount of viral mRNA in these cells. Figure 3C shows that it is a small fraction in many cells, but we can also state this fact in the text rather than just referring to the figure.

Overall I agree that textual changes will address reviewer misconceptions. We could add back in the nonstimulatory deletion I tested…but that does invite criticism given that it was identified in another, as-yet unpublished effort.

I also would be perfectly happy publishing in cell host & microbe resources, one of my PhD papers was published as the same and I have had no issues of impact from that paper, and indeed it is relatively well-cited.