Transcriptional dynamics of influenza virus infection in single cells

Alistair B. Russell[1], **Cole Trapnell**[2], **Jesse D. Bloom**[1,2]*

*For correspondence:
jbloom@fredhutch.org

[1]Basic Sciences Division and Computational Biology Program, Fred Hutchinson Cancer Research Center, Seattle, United States; [2]Department of Genome Sciences, University of Washington, Seattle, United States

Abstract   Influenza virus infection induces large changes in cellular transcription. Previously this has mostly been looked at using bulk measurements Here we examine the process at the level of single cells. We find extremely wide variation in the extent of viral gene transcription across infected cells. IFN induction is very rare. Some cellular pathways may be consistently altered in cells with high burden of viral transcripts. Overall, highlights remarkable heterogeneity in the outcome of infection.

Introduction

Heterogeneity is important in a lot of cellular processes even when isogenic (*Shalek et al., 2013*, *2014*).

Population (genetic) heterogeneity is also important. Viral quasispecies, cancer single-cell, etc. Salmonella paper (PhoP).

Literature on viral burst-size heterogeneity. This goes back to Delbruck, Andino polio paper (*Schulte and Andino, 2014*), the MDCK / flu paper.

Discuss segmented nature of influenza. Maybe in the context of how this could further increase heterogeneity because there is a lot of potential for entire genes to be missing. Includes Yewdell and Lowen papers.

Results

Strategy to measure mRNA transcription in single infected cells.

We chose to perform single-cell mRNA sequencing using a droplet-based system that physically isolates single cells prior to performing reverse-transcription. Each droplet contains a unique *cell barcode* that is used to tag all mRNAs from that cell during reverse-transcription. Each mRNA molecule is additionally tagged with a *unique molecular identifier (UMI)* during reverse transcription. Sequencing of the 3′ ends of the mRNAs therefore enables quantification of the number of different molecules of each mRNA species that are captured per cell.

Virally infected cells will express viral as well as cellular mRNAs. However, the cell barcodes and UMIs do not enable direct detection of whether a cell was initially infected by one or multiple viral particles. We therefore engineered an influenza virus that additionally carried *viral barcodes* near the 3′ end of each transcript (Figure 1A). These viral barcodes consisted of two synonymous mutations designed to fall in the 3′ portion of the mRNAs sequenced by our approach. These barcodes were engineered into reverse-genetics plasmids encoding the A/WSN/1933 (H1N1) strain
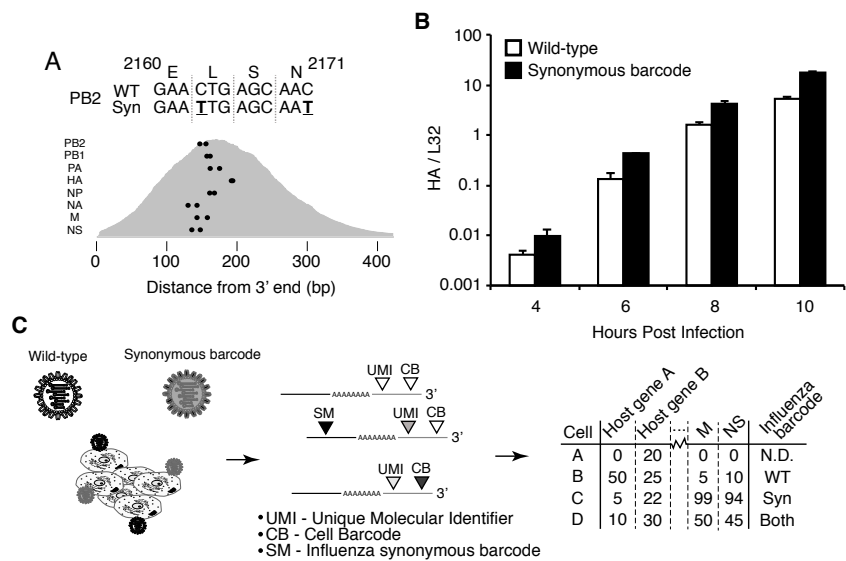
Figure 1. Possible changes: **(A)** Put current panel B first, since we want to explain how the viruses are barcoded before describing the infections. Also, perhaps show an example barcode (e.g., `GGACAG→GGcCAt`). **(B)** The wildtype and synonymously barcoded variants grow similarly. This is the current Figure 2A. **(C)** This would now be the current (A). We have plenty of space, so simply write out "wildtype", "synonymously barcoded", "cell barcode" even if it makes the figure a bit larger – this will be much clearer than introducing a bunch of acronyms. Also, I think the matrix could be made clearer. Add ellipses indicating that there are lots of genes and just a few are being shown. Also, make it "host gene A", "host gene B", ... For the viruses, I think it would be clearer to have columns "viral gene A, wildtype", "viral gene B, synonymous barcode", etc. That way you can have numbers in those columns which is what is actually present in the matrix.

Figure 1–source data 1. Sequences of wildtype and synonymously barcoded viruses are in FASTA file XXX (add this FASTA file to the figures directory).

of influenza virus, which were then used to generate virus. We also generated wildtype virus of the same strain in parallel, and validated that the two viral strains exhibited similar growth kinetics (Figure 1B). We then infected A549 human lung carcinoma cells with an equal mix of the wildtype and synonymously barcoded viruses. Cells infected by a single virion should exclusively express mRNAs from either wildtype or synonymously barcoded virus, while cells that are co-infected with multiple virions will frequently express mRNAs from both the wildtype and synonymously barcoded viruses (Figure 1C).

We also took great care to produce stocks of virus that we expected to be relatively "pure." Many viruses, including influenza, rapidly diversify into an array of biologically active viral particles, many of which can have mutations or deletions in essential viral genes. For influenza virus, one abundant forms of these defective particles arises from deletions that occur in the three viral segments encoding the tripartite viral polymerase. These defective particles become particularly abundant when the virus is grown at high multiplicities of infection, where virions with defective genes can be complemented by co-infection with other virions. To keep the level of these defective particles as low as possible, we propagated our viral populations at low multiplicity of infection. We then validated that compared to virus propagated at high multiplicity of infections, our viral populations exhibited a much greater purity of infectious particles. We validated this by assessing the ratio of infectious particles to viral-RNA molecules (Figure 2A) and to particles that induce cellular expression of a single viral protein (Figure 2B). We also showed that our viral populations induced much less interferon than a population propagated at higher multiplicity of infection as expected because... Lastly, our populations induced far less interferon than a high defective control, consistent with prior observations that such stocks are more immunostimulatory (Figure 2D).
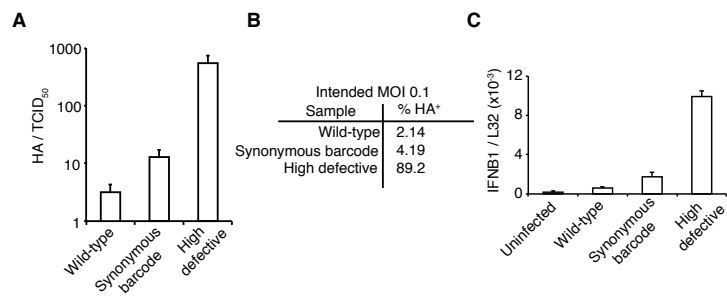
Figure 2. Panel A has now moved to the other figure. In panel A, the vertical alignment is a bit off for a few of the tick labels. Space between "HA" and "TCID50" on the y-axis.

Population-Level Measurements Validate Dataset.

Using these validated populations, we tracked growth kinetics in our target cell line at an infectious dose targeting a final infection rate of 5 percent as determined by flow cytometry. We observed rapid transcript accumulation through 10 hours, which then tapered off — potentially representing new rounds of infection rather than progression in initially infected cells (supplemental figure). We then chose three time points (6,8, and 10h) reflecting dramatically different influenza burden (and thus different stages in influenza replication) for this study.

Before proceeding with more detailed analyses, we first examined population-level data to confirm the validity of our approach. We not only generated data for our three time points and a matched uninfected control, but also performed a biological replicate at 8 hours to determine whether our findings could be dominated by inter-sample or technical variation rather than biological processes driven by influenza. We sampled roughly equivalent number of cells per time point at an equivalent sequencing depth per cell - granting similar probability to observe rare events across datasets (Figure 3A) (ADD SUPPLEMENTAL OF DEPTH/CELL?). Overall we observed highly similar distributions of transcript abundance between samples (Figure 3B). This suggests that, in a rough global sense, transcriptome dynamics are relatively unperturbed and that influenza, if it is driving dramatic changes in transcript abundance and/or growth state, must be doing so in only a small subset of cells. More importantly, with the confidence that the average cellular state is similar between time points, we can better regard changes in subpopulations between time points as bone-fide progression in influenza replication rather than inter-sample variation.

In addition to analyses regarding global transcript abundance, we were also able to successfully detect increasing numbers of influenza transcripts as our experiment progressed (Figure 3C). Moreover, there appeared to be only a few cells at each time point making considerable contributions to the influenza transcript pool, consistent with our desired low level of infection. With these metrics, we felt confident proceeding to more detailed analyses of influenza replication kinetics and impact on these populations.

Synonymous Barcodes Provide An Empirical Means of Setting Infection Thresholds.

As a first step in understanding influenza replication at a single-cell level, we must derive a threshold for identifying infected cells. Crucial to our experimental design, we were able to detect both wild-type and synonymous influenza variants at all three timepoints in all eight segments, at levels permitting identification of cells infected with one, or both strains (Figure 4-Figure supplement 1). This was true of about half of the transcripts surveyed, matching the read-density of our preliminary experiments. These barcodes partitioned non-randomly between cells, as expected if derived from single infectious events (Figure 4A,B). This grants us the power to develop an empirical threshold for calling influenza-infected cells versus those that might have reads derived from lysis
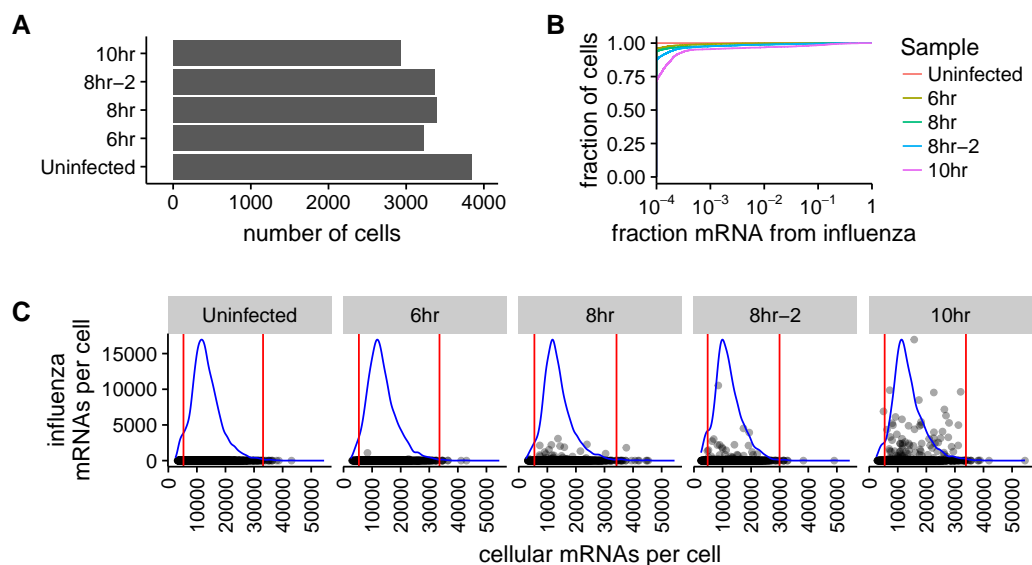
Figure 3. Overview of amounts of cellular and influenza virus mRNAs detected in each cell. **(A)** Number of cells captured for each sample. **(B)** Cumulative fraction plot showing the amount of mRNA derived from influenza for each sample. In all samples, most cells had little or no influenza mRNA. **(C)** The number of cellular and viral mRNAs for each cell is plotted as a point. The blue lines show the overall distribution of the number of cellular mRNAs per sample. Cells that fell outside the red lines were removed as outliers. At later timepoints, a small number of cells had a very high number of viral mRNAs.

⁹⁷ of infected cells. Reads derived from lysed cells should exhibit random partitioning ? sampling each
⁹⁸ barcode according to its overall abundance in the experiment. Therefore, when ordering all cells at
⁹⁹ all timepoints according to the number of influenza reads observed and the relative abundance of
¹⁰⁰ each barcode, we can set our detection threshold as the point at which barcodes reliably exhibit
¹⁰¹ non-random partitioning - that is when all segments in a given cell are frequently derived from
¹⁰² either our synonymous or wild-type virus but not both (Figure 4C). Notably, this strategy relies
¹⁰³ heavily on a relatively low MOI, as high rates of coinfection would stymie this approach (Figure 4D,E)
¹⁰⁴ .
¹⁰⁵ Few coinfections
¹⁰⁶ Setting threshold: Assumes rare barcode events are due to lysis (or misannotation) and that
¹⁰⁷ two values SHOULD correlate Validating evidence/assumption two values correlate - flow data, flow
¹⁰⁸ diagram and quantile-normalized data. Include replicates in supplement.
¹⁰⁹ Once defined, we have some statistics.
¹¹⁰ Presumably this subsection basically covers 4.

Influenza Exhibits A Broad Distribution of Productivity Partially Explained By Segment Absence.

¹¹³ The absolute productivity of influenza infection as measured by the number of fully-infectious
¹¹⁴ virions produced by a given cell is known to vary dramatically. This observation is recapitulated
¹¹⁵ in our data, wherein of the small portion of infected cells, a bare few contribute the majority of
¹¹⁶ influenza transcripts (Figure 4-Figure supplement 2). As transcript abundance is highly dependent
¹¹⁷ upon secondary transcript from de novo generated viral genomes, it can be presumed that when
¹¹⁸ comparing between two cells, increased mRNA abundance is likely correlated with increased
¹¹⁹ viral genome production, which in turn is highly predictive of viral productivity. It is thought that
¹²⁰ one driver of the observed variation in influenza productivity is the segmented nature of the
¹²¹ influenza genome. Several groups have described that the majority of infections fail to express at
¹²² least one of the eight influenza segments. Our datasets produce a similar pattern, even in cells
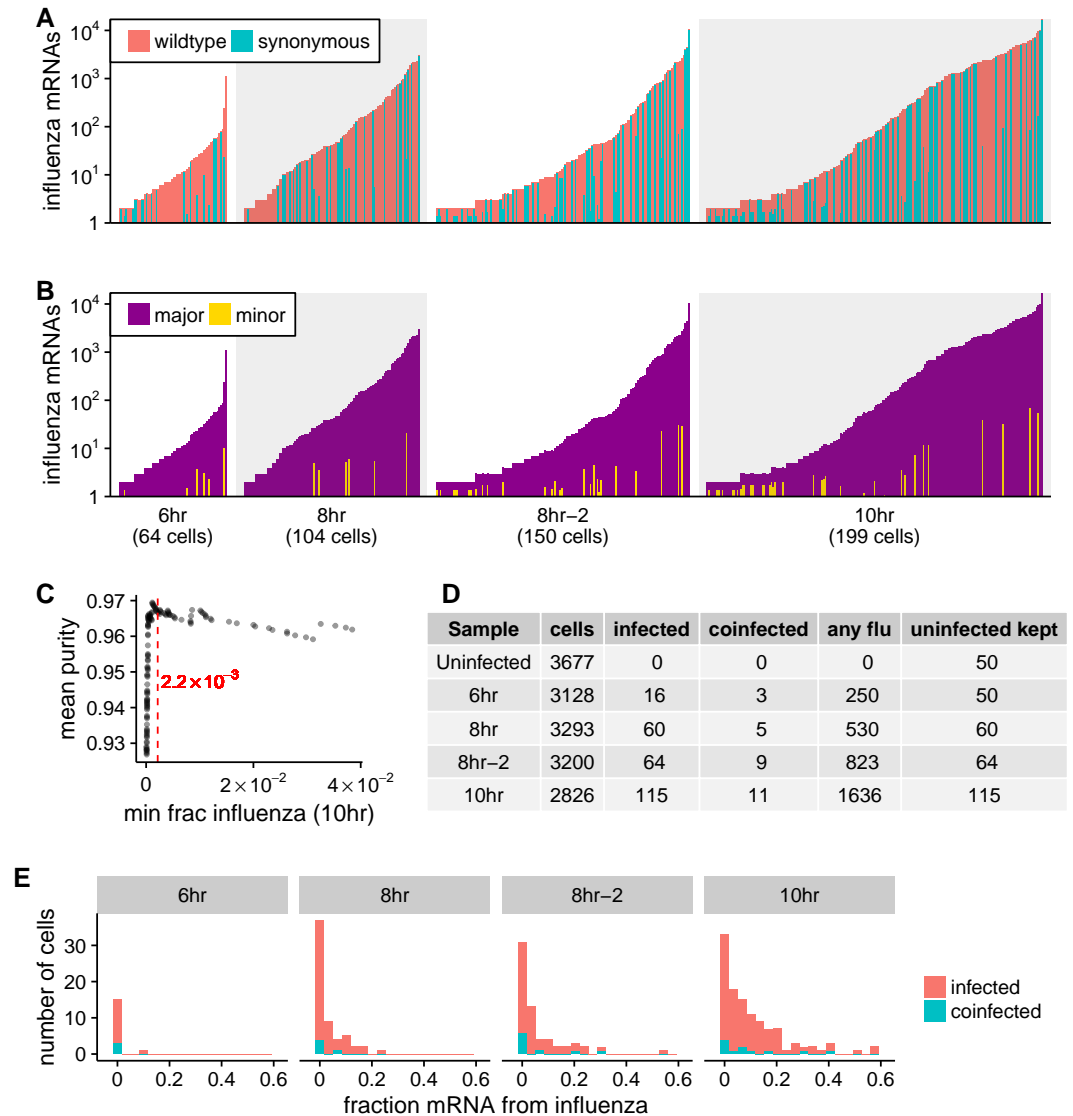
Figure 4. Synonymous barcodes near the 3' end of the influenza virus mRNAs were used to identify co-infection and distinguish true infections from cells that contained a few spurious viral reads. **(A)** For all cells with at least two viral mRNAs for which the synonymous barcode could be called, each line is proportional to the logarithm of the number of viral mRNAs in that cell. The bars are colored in linear proportion to the fraction of the viral mRNAs derived from either wildtype or synonymously barcoded virus. **(B)** Same as (A), but now each bar is colored according to the relative proportions of the more common (major) and less common (minor) barcoded virus variant. At low levels of viral mRNA there is often a roughly equal mix of barcodes, since many of these cells have simply picked up environment mRNA which is about equally likely to derive from either virus. But at higher levels of viral mRNA, truly infected cells are mostly one pure barcode except for a few cells that are truly co-infected. **(C)** We determined a cutoff for calling "true" infections by fitting a curve to the mean barcode purity of all cells with greater than a given fraction of their mRNA derived from virus. We called the cutoff at the point at which purity stops increasing with the fraction of viral mRNA. **(D)** The number of cells identified as infected and co-infected for each sample. For all samples, the vast majority of cells were not infected, so for subsequent analyses we subsampled to a number of uninfected cells that was the greater of 50 or the number of infected cells. **(E)** The distribution of the fraction of mRNA derived from virus for each sample for both infected and co-infected cells. For all samples, there is a very wide distribution of the amount of viral mRNA.
Figure 4–Figure supplement 1. Number of viral barcodes called.
Figure 4–Figure supplement 2. Total fraction of viral mRNA attributable to a given fraction of infected cells.

123 that are otherwise expressing high levels of influenza transcripts (Figure 5A). When we bin cells
124 according to absence of a given segment we observe that, unsurprisingly, absence of the tripartite
125 polymerase (PB1, PA, PB2) and nuceoprotein (NP) correlates strongly with decreased influenza
126 burden (Figure 5B).. Importantly, the gene products of all four of these segments are essential to
127 the transition from primary mRNA transcription to replication and secondary mRNA transcription.
128 This effect cannot be solely attributed to sampling error, as the polymerase genes exhibit the lowest
129 expression and thus would be expected to be stochastically lost due to undersurveillance at much
130 higher influenza transcript levels than would other genes.

131 Nevertheless, while segment absence can partially explain extreme components of influenza
132 variability resulting from an inability to fully engage in secondary transcription, absence of the re-
133 maining segments does not appear to have any significant impact on influenza burden. Importantly,
134 even cells expressing all eight segments still exhibit a wide distribution of influenza transcripts
135 within each time point. We must instead therefore posit alternative mechanisms such as variability
136 in host cells to explain the remainder of this distribution. Our data do differ slightly on this point
137 than prior observations by Brooke et al, inasmuch as their report described decreased influenza
138 burden not only associated with loss of nucleoprotein, but also of the NS segment. However,
139 significant strain-to-strain differences exist in the behavior of NS1, a major interferon antagonist
140 encoded by the NS segment, and these differences might explain this discrepancy.

141 Extrapolating these measurments to?.

Influenza Infections Are Predominated By Nonproductive Events.

143 We further wished to define metrics for presence/absence for a given segment and compare
144 them to values obtained for other strains of influenza in order to understand the probability that a
145 given infection could possibly produce infectious progeny. We decided to omit these calculations for
146 PB1, PA, PB2, and NP. Using cyclohexamide treatment to limit cells to primary transcription alone,
147 we find a 2-3 log reduction in mRNA levels of PB1 or HA relative to a 10h time point (supplemental
148 figure). We therefore are not confident that our approach has the power to detect all semi-infectious
149 events that result from a lack of polymerase or nucleoprotein segments. However, given the lack of
150 effect absence of other segments appears to have on mRNA production, we proceeded to calculate
151 the relative abundance of infectious events absent a given gene segment.

Influenza mRNA Ratios are Tightly Controlled

153 This and other reports have described the considerable heterogeneity produced as influenza
154 infection progresses in tissue culture models. However, in stark contrast to our analyses above,
155 the ratios of mRNAs generated by the eight segments appear to behave in a highly stereotyped
156 fashion, with values roughly consistent with prior Northern Blot measurements of Influenza mRNA
157 abundance. Staggeringly, this is true whether or not cells are expressing mRNAs from all eight
158 segments, at all three time points, and, in cells running the gamut of infectious burden — although
159 we must limit these conclusions to those cells engaging in secondary transcription given the
160 caveats described above. Therefore while either noise in influenza replication or host variability can
161 dramatically impact the amount of Influenza mRNA produced as a whole, these processes seem to,
162 ignoring complete absence of a segment, act equally across the Influenza genome. This is consistent
163 with more limited measurements showing positive correlations between several segments in vRNA
164 measurements in single cells. It would therefore be inadvisable to search for the sources of the
165 broad distribution in influenza replication in those steps that can impact segments independently
166 of one-another, such as noise in the transition to cRNA or vRNA production, but rather only in
167 those processes that could impact the kinetics of all segments equivalently, such as transport to
168 the nucleus or availability of cellular resources.

A

fraction mRNA from flu: • lacks RNP gene • lacks other gene • all 8 genes
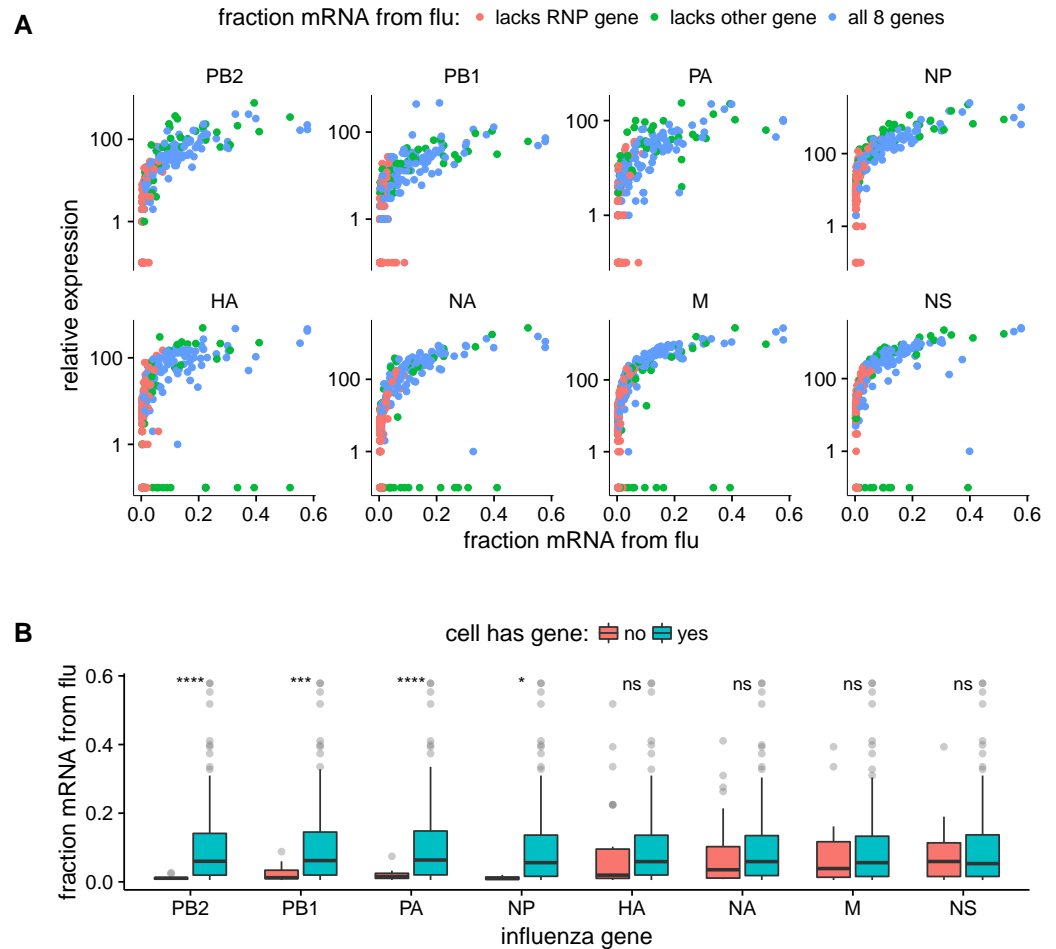

B

cell has gene: □ no □ yes

Figure 5. The viral infection burden in individual cells as a function of the amount of each viral gene detected.
(A) Fraction of mRNAs in each cell derived from virus as a function of the *normalized* expression of each viral gene in that cell. This plot shows that all cells with very high viral burden express all of the RNP genes, but some cells with high viral burden lack each of the other four viral genes. **(B)** Statistical tests confirming that absence of viral RNP genes is significantly associated with reduced viral burden, but that the absence of the non-RNP genes does not lead to a clear decrease in viral burden.
Figure 5–Figure supplement 1. Like panel (B) but for the 10-hr sample only.
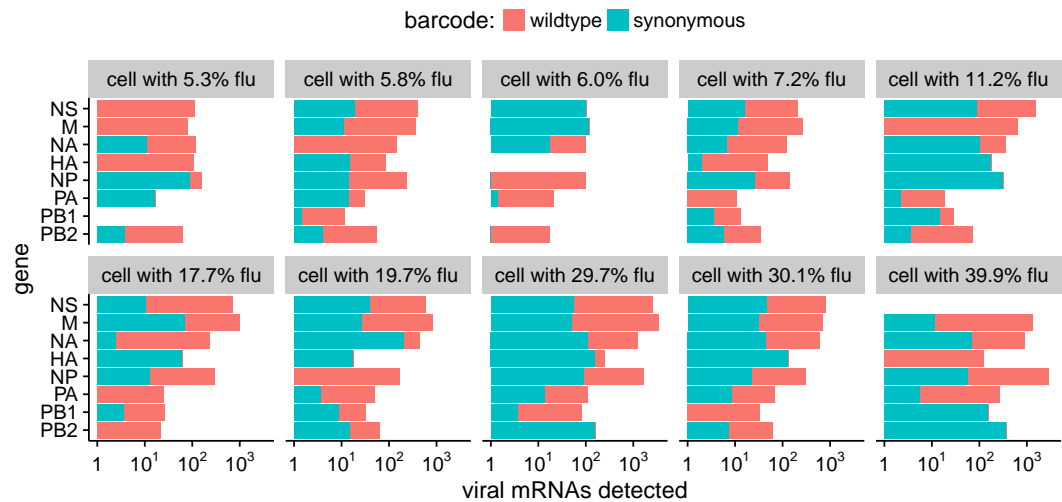
Figure 6. Frequency of each viral gene segment in co-infected cells with at least 5% of their mRNA derived from influenza. The bars indicate the logarithm of the number of each viral mRNA detected, and the bars are colored in proportion to the fraction of those mRNAs that are derived from either wildtype or synonymous barcoded virus.

Figure 6–source data 1. The raw data plotted in this figure are in `p_coinfection.csv`.

Coinfection Occurs Early and Is Likely Adaptive For Influenza

169 While relatively rare in our dataset, those cells positively identified as containing both wild-type
170 and synonymously barcoded influenza provide important insight into the nature of coinfection. It
171 has been posited that coinfection can rescue nonproductive infections, potentially by providing
172 missing genome equivalents. Brooke et al., for example observed increased co-occurrence of a
173 given gene-segment pair as MOI was increased. When we examine our coinfected cells, we observe
174 a much higher rate of expression of all eight segments (metric). Consistent with this model, we
175 also observe that if we consider only one barcode or the other, that many of these infections
176 would instead have been non-productive. Another observation that can be immediately procured
177 is that, on the whole, coinfections have roughly balanced amounts of wild-type and barcoded
178 segments. Others have noted that windows for coinfection are relatively narrow, and, these data
179 directly suggest that this window is specifically narrow with respect to cRNA and vRNA production.
180 If significant replication of viral genomes had already occurred, one would expect a much wider
181 distribution of ratios than those observed. To support these observations, we used a virus in which
182 the HA segment was replaced by GFP, allowing us to detect co-infection by this modified virus and
183 wild-type WSN by flow cytometry. When we examine strictly co-infected cells, we observe that, as
184 in our single-cell analysis, both HA and GFP are highly correlated — indicating a tight window of
185 coinfection — and that a peak of low HA — likely corresponding to primary transcription alone
186 — is greatly reduced. These data give credence to the sentiments of others that non-productive
187 events are perhaps just productive events waiting to occur, and that the true dynamics underlying
188 descriptions of ?MOI? are far more complex than we frequently give them credence and rely on
189 assumptions that frequently prove incorrect.

Relative expression of different viral genes

190 Figure 7 shows some of this. Cite *Hatada et al.* (*1989*) to show our relative expressions are
191 consistent with that. Maybe show a panel with just the 8-segment infections.

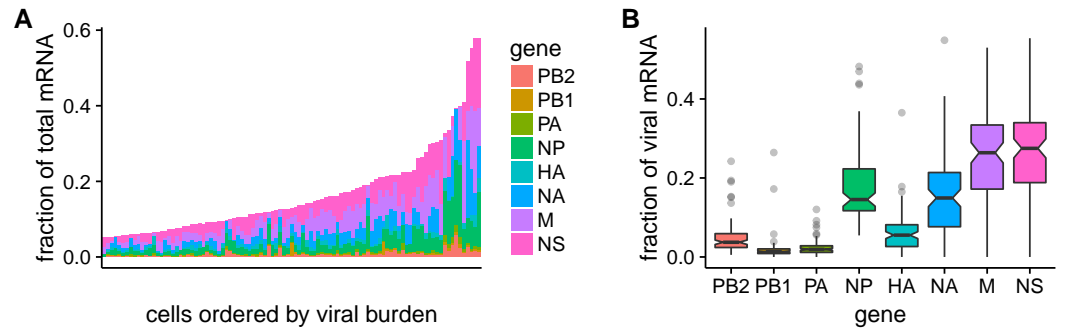192 Maybe add a figure breaking this down among timepoints.

Host stuff

Figure 7. Expression of individual influenza genes in highly infected cells (at least 5% of total mRNA is viral). **(A)** The fraction of total mRNA from each influenza gene for each cell. **(B)** Box plots showing the fraction of viral mRNA per cell that is derived from each influenza gene taken over all highly expressed cells. The black line at the notch in each box is the median, and the top and bottom of the box indicate the first and third quartiles.

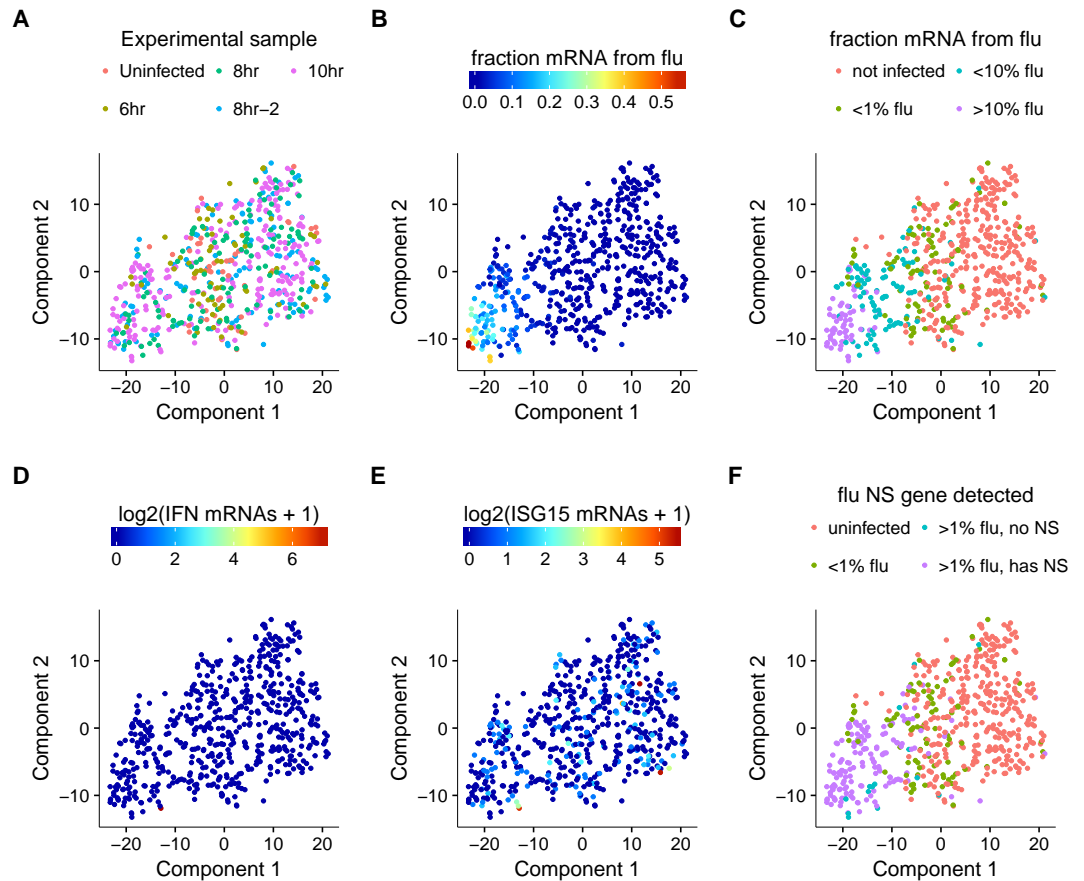**Figure 7–source data 1.** The raw data are in `p_flu_expr.csv`.

Figure 8. tSNE plots. The layout is the same in all panels, but each panel colors the cells according to a different property. **(A), (B)** Cells colored by the fraction of their mRNA that is viral. **(C)** Cells colored by experimental sample. While it is clear that cells from later timepoints often have more viral RNA, there are cells from earlier timepoints with a high viral burden and cells from late timepoints with a low viral burden. **(D)** Cells colored by the number of type I and III interferon transcripts detected. Only one cell has high expression of these interferons. **(E)** Cells colored by the expression of the interferon-stimulate gene ISG15. **(F)** For cells with at least 1% of their reads from influenza, are the cells expressing the viral NS protein? The one interferon-positive cell is lacking NS, but many other cells also lack NS but do not express interferon.
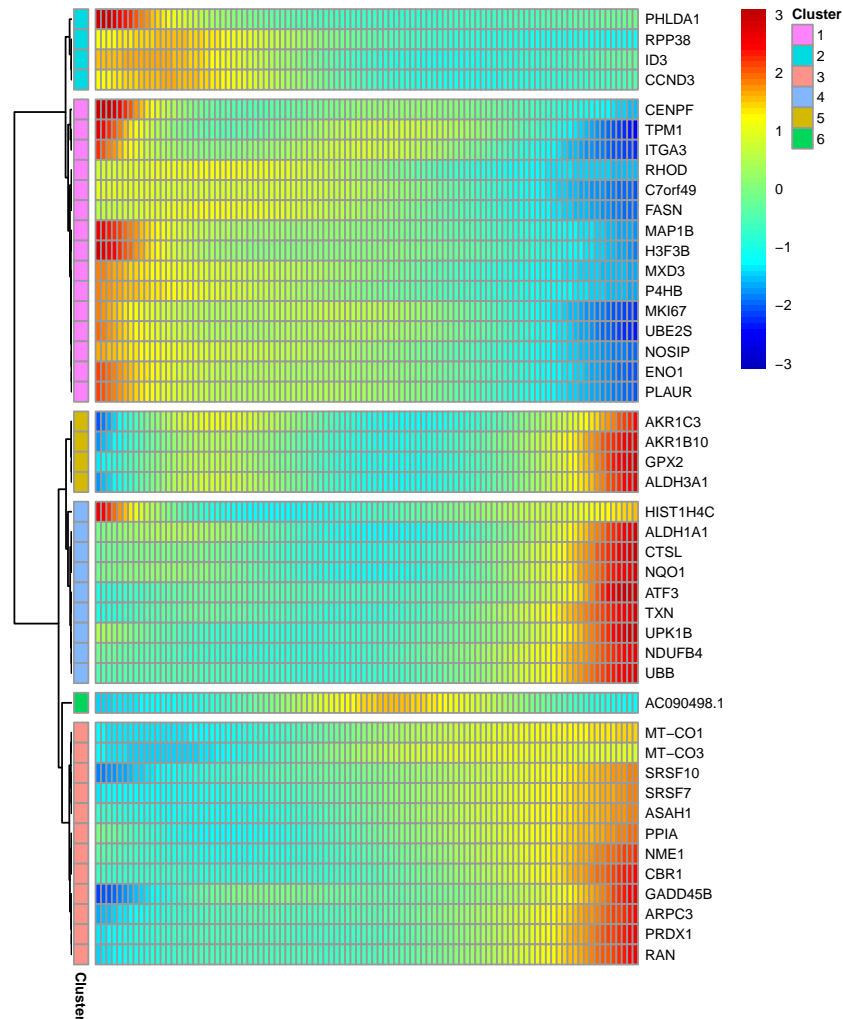
Figure 9. Cellular genes that are differentially expressed with respect to the amount of influenza mRNA in individual cells infected with full influenza virus containing all eight genes. Shown are all genes differentially expressed with $Q < 0.1$.

Figure 9–source data 1. The full results of the differential expression test is in `p_sig_cellular_genes.csv`.

Figure 9–source data 2. The results of a gene-set analysis are in `p_sig_cellular_genes.csv`.

Discussion

Morbi luctus, wisi viverra faucibus pretium, nibh est placerat odio, nec commodo wisi enim eget quam. Quisque libero justo, consectetuer a, feugiat vitae, porttitor eu, libero. Suspendisse sed mauris vitae elit sollicitudin malesuada. Maecenas ultricies eros sit amet ante. Ut venenatis velit. Maecenas sed mi eget dui varius euismod. Phasellus aliquet volutpat odio. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Pellentesque sit amet pede ac sem eleifend consectetuer. Nullam elementum, urna vel imperdiet sodales, elit ipsum pharetra ligula, ac pretium ante justo a nulla. Curabitur tristique arcu eu metus. Vestibulum lectus. Proin mauris. Proin eu nunc eu urna hendrerit faucibus. Aliquam auctor, pede consequat laoreet varius, eros tellus scelerisque quam, pellentesque hendrerit ipsum dolor sed augue. Nulla nec lacus.

Methods and Materials

Guidelines can be included for standard research article sections, such as this one.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Some LaTeX Examples

Use section and subsection commands to organize your document. LaTeX handles all the formatting and numbering automatically. Use ref and label commands for cross-references.

Figures and Tables

If you use the following prefixes for your `\label`:

Figures `fig:`, e.g. `\label{fig:view}`

Tables `tab:`, e.g. `\label{tab:example}`

Equations `eq:`, e.g. `\label{eq:CLT}`

Boxes `box:`, e.g. `\label{box:simple}`

you can then use the convenience commands as in `\FIG{cells}`, to generate cross-reference **??**.

Citations

LaTeX formats citations and references automatically using the bibliography records in your .bib file, which you can edit via the project menu. Use the `\cite` command for an inline citation, like *Trapnell et al.* (*2014*), and the `\citep` command for a citation in parentheses (*Trapnell et al., 2014*). The LaTeX template uses a slightly-modified Vancouver bibliography style. If your manuscript is accepted, the eLife production team will re-format the references into the final published form. *It is not necessary to attempt to format the reference list yourself to mirror the final published form.*

Acknowledgments

Additional information can be given in the template, such as to not include funder information in the acknowledgments section.

References

Akpinar F, Timm A, Yin J. High-throughput single-cell kinetics of virus infections in the presence of defective interfering particles. Journal of virology. 2016; 90(3):1599–1612.

Baum A, Sachidanandam R, García-Sastre A. Preference of RIG-I for short viral RNA molecules in infected cells revealed by next-generation sequencing. Proceedings of the National Academy of Sciences. 2010; 107(37):16303–16308.

Bhushal S, Wolfsmüller M, Selvakumar TA, Kemper L, Wirth D, Hornef MW, Hauser H, Köster M. cell Polarization and epigenetic status shape the heterogeneous response to Type iii interferons in intestinal epithelial cells. Frontiers in Immunology. 2017; 8.

Brooke CB, Ince WL, Wei J, Bennink JR, Yewdell JW. Influenza A virus nucleoprotein selectively decreases neuraminidase gene-segment packaging while enhancing viral fitness and transmissibility. Proceedings of the National Academy of Sciences. 2014; 111(47):16854–16859.

Brooke CB, Ince WL, Wrammert J, Ahmed R, Wilson PC, Bennink JR, Yewdell JW. Most influenza a virions fail to express at least one essential viral protein. Journal of virology. 2013; 87(6):3155–3162.

Combe M, Garijo R, Geller R, Cuevas JM, Sanjuán R. Single-cell analysis of RNA virus infection identifies multiple genetically diverse viral genomes within single infectious units. Cell host & microbe. 2015; 18(4):424–432.

Davis AR, Hiti AL, Nayak DP. Influenza defective interfering viral RNA is formed by internal deletion of genomic RNA. Proceedings of the National Academy of Sciences. 1980; 77(1):215–219.

Delbrück M. The burst size distribution in the growth of bacterial viruses (bacteriophages). Journal of bacteriology. 1945; 50(2):131.

Dou D, Hernández-Neuta I, Wang H, Östbye H, Qian X, Thiele S, Resa-Infante P, Kouassi NM, Sender V, Hentrich K, et al. Analysis of IAV Replication and Co-infection Dynamics by a Versatile RNA Viral Genome Labeling Method. Cell Reports. 2017; 20(1):251–263.

Hatada E, Hasegawa M, Mukaigawa J, Shimizu K, Fukuda R. Control of influenza virus gene expression: quantitative analysis of each viral RNA species in infected cells. The Journal of Biochemistry. 1989; 105(4):537–546.

Heldt FS, Kupke SY, Dorl S, Reichl U, Frensing T. Single-cell analysis and stochastic modelling unveil large cell-to-cell variability in influenza A virus infection. Nature communications. 2015; 6:8938.

Hutchinson EC, Curran MD, Read EK, Gog JR, Digard P. Mutational analysis of cis-acting RNA signals in segment 7 of influenza A virus. Journal of virology. 2008; 82(23):11869–11879.

Kawakami E, Watanabe T, Fujii K, Goto H, Watanabe S, Noda T, Kawaoka Y. Strand-specific real-time RT-PCR for distinguishing influenza vRNA, cRNA, and mRNA. Journal of virological methods. 2011; 173(1):1–6.

Killip MJ, Fodor E, Randall RE. Influenza virus activation of the interferon system. Virus research. 2015; 209:11–22.

Killip MJ, Jackson D, Pérez-Cidoncha M, Fodor E, Randall RE. Single-cell studies of IFN-β promoter activation by wild-type and NS1-defective influenza A viruses. Journal of General Virology. 2017; 98(3):357–363.

López CB. Defective viral genomes: critical danger signals of viral infections. Journal of virology. 2014; 88(16):8720–8723.

Odagiri T, Tashiro M. Segment-specific noncoding sequences of the influenza virus genome RNA are involved in the specific competition between defective interfering RNA and its progenitor RNA segment at the virion assembly step. Journal of virology. 1997; 71(3):2138–2145.

Pérez-Cidoncha M, Killip MJ, Oliveros JC, Asensio VJ, Fernández Y, Bengoechea JA, Randall RE, Ortín J. An unbiased genetic screen reveals the polygenic nature of the influenza virus anti-interferon response. Journal of virology. 2014; 88(9):4632–4646.

Saira K, Lin X, DePasse JV, Halpin R, Twaddle A, Stockwell T, Angus B, Cozzi-Lepri A, Delfino M, Dugan V, et al. Sequence analysis of in vivo defective interfering-like RNA of influenza A H1N1 pandemic virus. Journal of virology. 2013; 87(14):8064–8074.

Schulte MB, Andino R. Single-cell analysis uncovers extensive biological noise in poliovirus replication. Journal of virology. 2014; 88(11):6205–6212.

Shalek AK, Satija R, Adiconis X, Gertner RS, Gaublomme JT, Raychowdhury R, Schwartz S, Yosef N, Malboeuf C, Lu D, et al. Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. Nature. 2013; 498(7453):236.

287 **Shalek AK**, Satija R, Shuga J, Trombetta JJ, Gennert D, Lu D, Chen P, Gertner RS, Gaublomme JT, Yosef N, et al.
288   Single cell RNA Seq reveals dynamic paracrine control of cellular variation. Nature. 2014; 510(7505):363.

289 **Shapiro G**, Gurney T, Krug R. Influenza virus gene expression: control mechanisms at early and late times of
290   infection and nuclear-cytoplasmic transport of virus-specific RNAs. Journal of virology. 1987; 61(3):764–773.

291 **Tapia K**, Kim Wk, Sun Y, Mercado-López X, Dunay E, Wise M, Adu M, López CB. Defective viral genomes arising
292   in vivo provide critical danger signals for the triggering of lung antiviral immunity. PLoS pathogens. 2013;
293   9(10):e1003703.

294 **Trapnell C**, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, Lennon NJ, Livak KJ, Mikkelsen TS, Rinn JL.
295   Pseudo-temporal ordering of individual cells reveals dynamics and regulators of cell fate decisions. Nature
296   Biotechnology. 2014; 32(4):381.

297 **Vreede FT**, Jung TE, Brownlee GG. Model suggesting that replication of influenza virus is regulated by stabilization
298   of replicative intermediates. Journal of virology. 2004; 78(17):9568–9572.

299 **Zhu Y**, Yongky A, Yin J. Growth of an RNA virus in single cells reveals a broad fitness distribution. Virology. 2009;
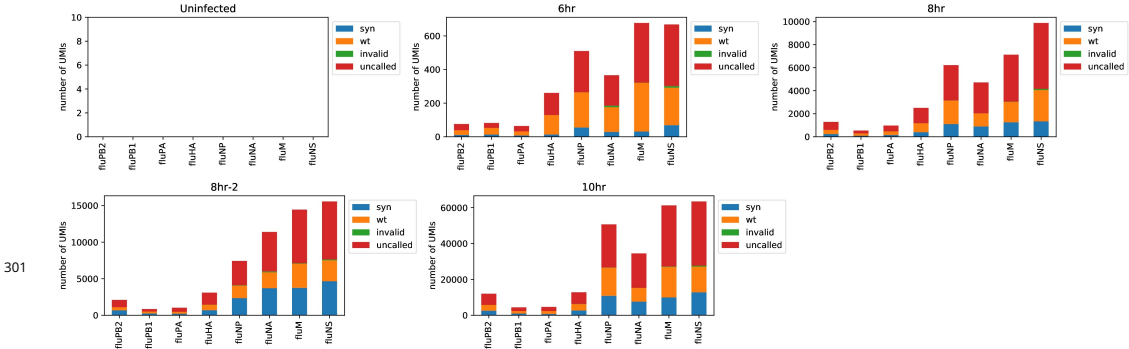300   385(1):39–46.

Figure 4–Figure supplement 1. The number of viral barcodes called for each sample and influenza gene segment. Viral transcripts are classified as *syn* if they mapped to a synonymously barcoded influenza transcript, *wt* if they mapped to a wildtype influenza transcript, *invalid* if multiple reads for the same UMI differed on the status of the viral barcode, and as *uncalled* if none of the reads for that UMI overlapped the region of the viral transcript containing the viral barcode.
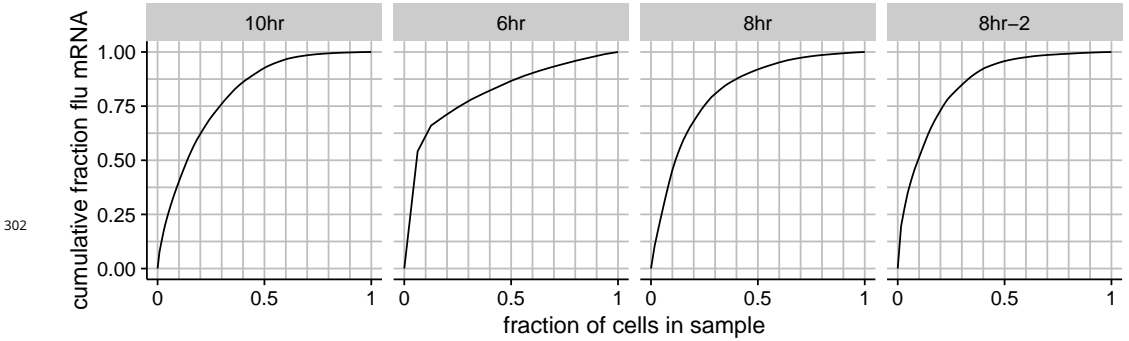

Figure 4–Figure supplement 2. The cumulative fraction of total viral mRNA among all infected cells that is attributable to a given fraction of cells. In all samples, more than half the total viral mRNA is derived from less than half the cells.
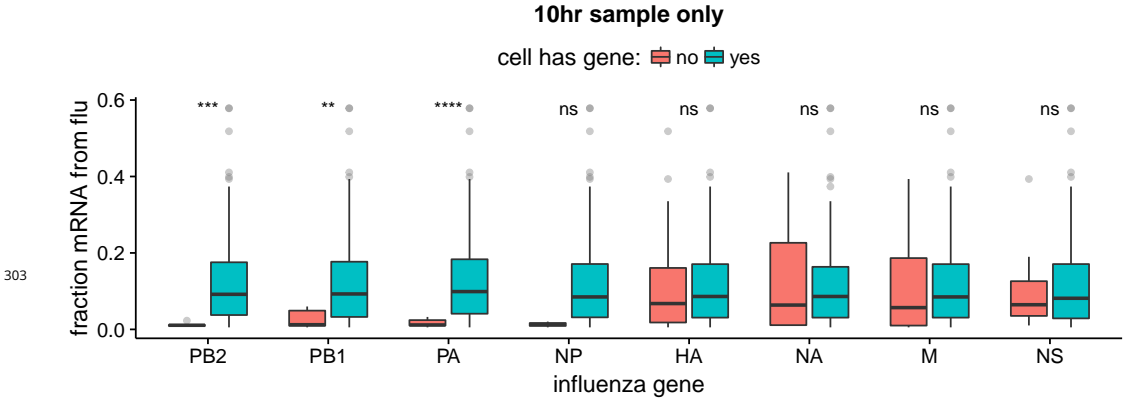

Figure 5–Figure supplement 1. The absence of viral RNP genes but *not* non-RNP genes remains significantly associated with reduced viral burden when we examine only the 10-hr sample, which is the single timepoint with the most data points. The difference for NP is no longer statistically significant due to low counts of infected cells lacking NP, but the qualitative trend remains.