Transcriptional dynamics of influenza virus infection at the single-cell level

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

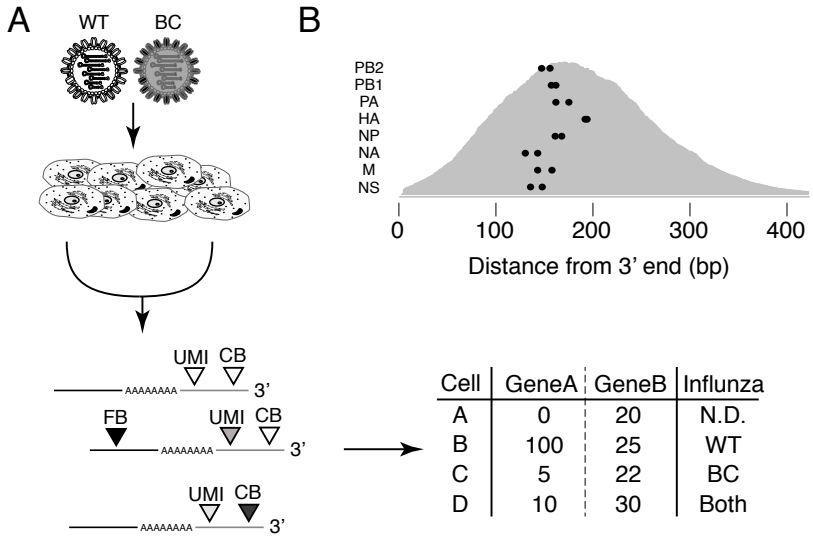
Generation of influenza virus stocks for single-cell infections.

Schematic of how the experiment works and generation of barcoded virus (Figure 1).

Validation of viral populations: wildtype and synonymous barcode (we need to pick some consistent nomenclature) grow about the same (Figure 2A). Maybe need a sentence or two about defectives should be working in as background before the next sentence (Carolyn Lopes paper on flu, some other stuff). The ratio of HA molecules to infectious virus particles is much lower than for a high-defective population (Figure 2B). These viruses don't induce very much IFN both overall and compared to a high-defective population (Figure 2C). The number of HA-expressing infection events is not all that much higher than the number of actual infectious events, in contrast to a high-defective population (Figure 2D).

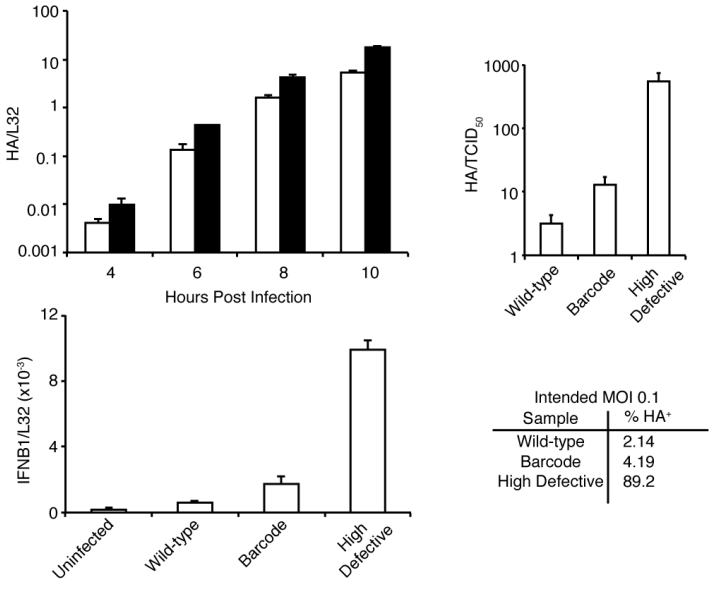
Global analysis of single-cell transcript counts reveals robustness of approach.

We got some reasonable number of cells at several timepoints (Figure 3A).



Cell	GeneA	GeneB	Influnza
A	0	20	N.D.
B	100	25	WT
C	5	22	BC
D	10	30	Both

Figure 1. CAPTION



Intended MOI 0.1

Sample	% HA⁺
Wild-type	2.14
Barcode	4.19
High Defective	89.2

Figure 2. CAPTION

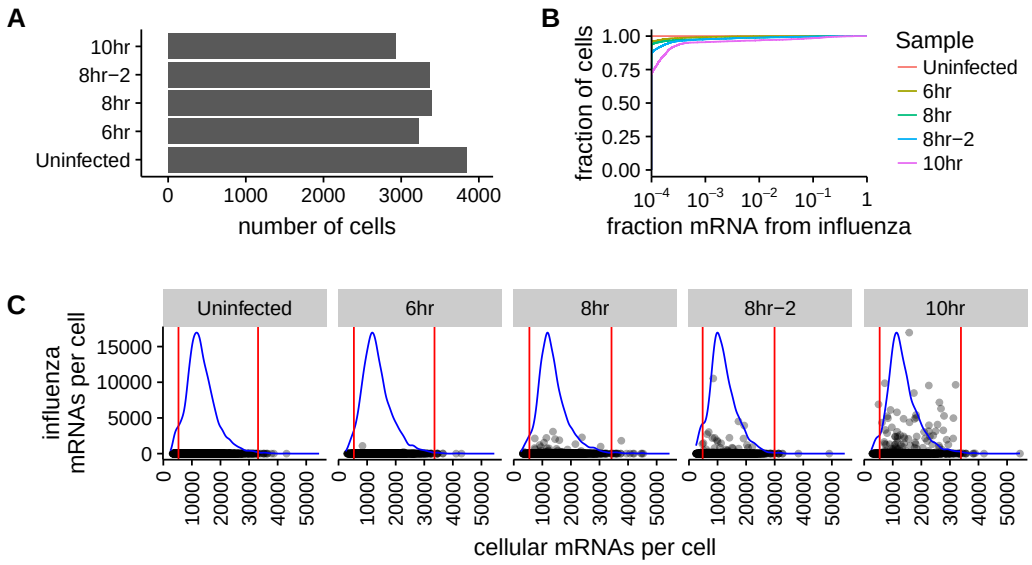


Figure 3. Overview of amounts of cellular and influenza virus mRNAs detected in each cell. **(A)** Number of cells captured for each sample. **(B)** Cumulative fraction plot showing the amount of mRNA derived from influenza for each sample. In all samples, most cells had little or no influenza mRNA. **(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.

Figure 3–Figure supplement 1. Shorter caption for main text.

Figure 3–Figure supplement 2. This is another supplementary figure.

Figure 3–source data 1. This is a description of a data source.

Most cells are uninfected, flu burden generally increases with time in the samples, show Figure 3B maybe with inset. Also show Figure 3C which basically shows the same thing in a different way.

We observe both barcodes in influenza-infected cells, and can determine the barcode for about half the viral transcripts (Figure 4).

Barcodes enable systematic identification of infected and co-infected cells.

Few coinfections

Setting threshold: Assumes rare barcode events are due to lysis (ormisannotation) and that two values SHOULD correlate Validating evidence/assumption two values correlate - flow data, flow diagram and quantile-normalized data. Include replicates in supplement.

Once defined, we have some statistics.

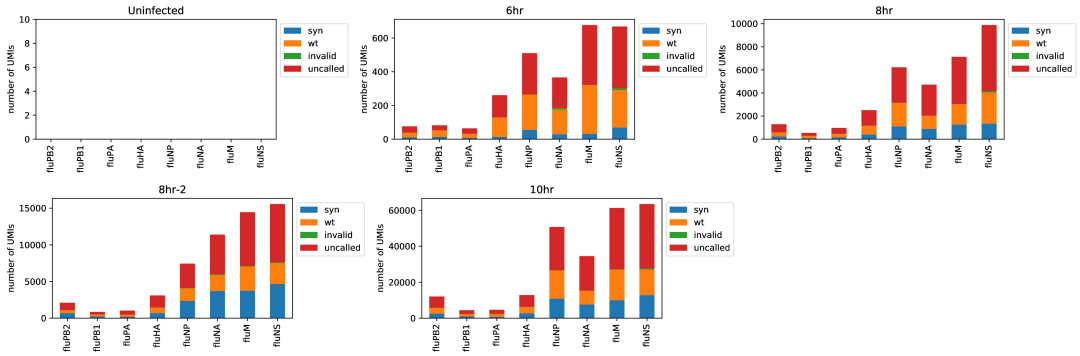


Figure 4. CAPTION

Presumably this subsection basically covers 5.

Absence of viral genes explains some of the heterogeneity in viral burden.

In the prior sections, we've pointed that there is a lot of variation in how much flu there is in each cell. One obvious explanation is some cells essential flu genes. Most obviously, viruses missing a polymerase gene are presumably have problems making transcripts. Possibly one could imagine that missing other genes would have an effect.

Figure **??** shows that viruses missing polymerase genes never get up as high in flu burden. Also work in as a figure supplement the qPCR of the cycloheximide.

Talk about caveats of calling polymerase absence.

Potentially add a figure or figure panel showing that there is still a lot of heterogeneity in 8-segment virus.

How many cells are missing influenza genes?

Look at the non-polymerase genes in highly infected cells, and see how often they are absent. This probably requires a new figure.

Talk about how co-infection can rescue absence of genes. From our single-cell data, Figure 7 shows some of the relevant data.

From the other data, talk about the flow cytometry data. Coinfection gives multiple 'chances' bring up flow data (hiDI and co-infection, as well as available data, bring up prior Yewdell data consistent with this).

Relative expression of different viral genes

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

Maybe add a figure breaking this down among timepoints.

Host stuff

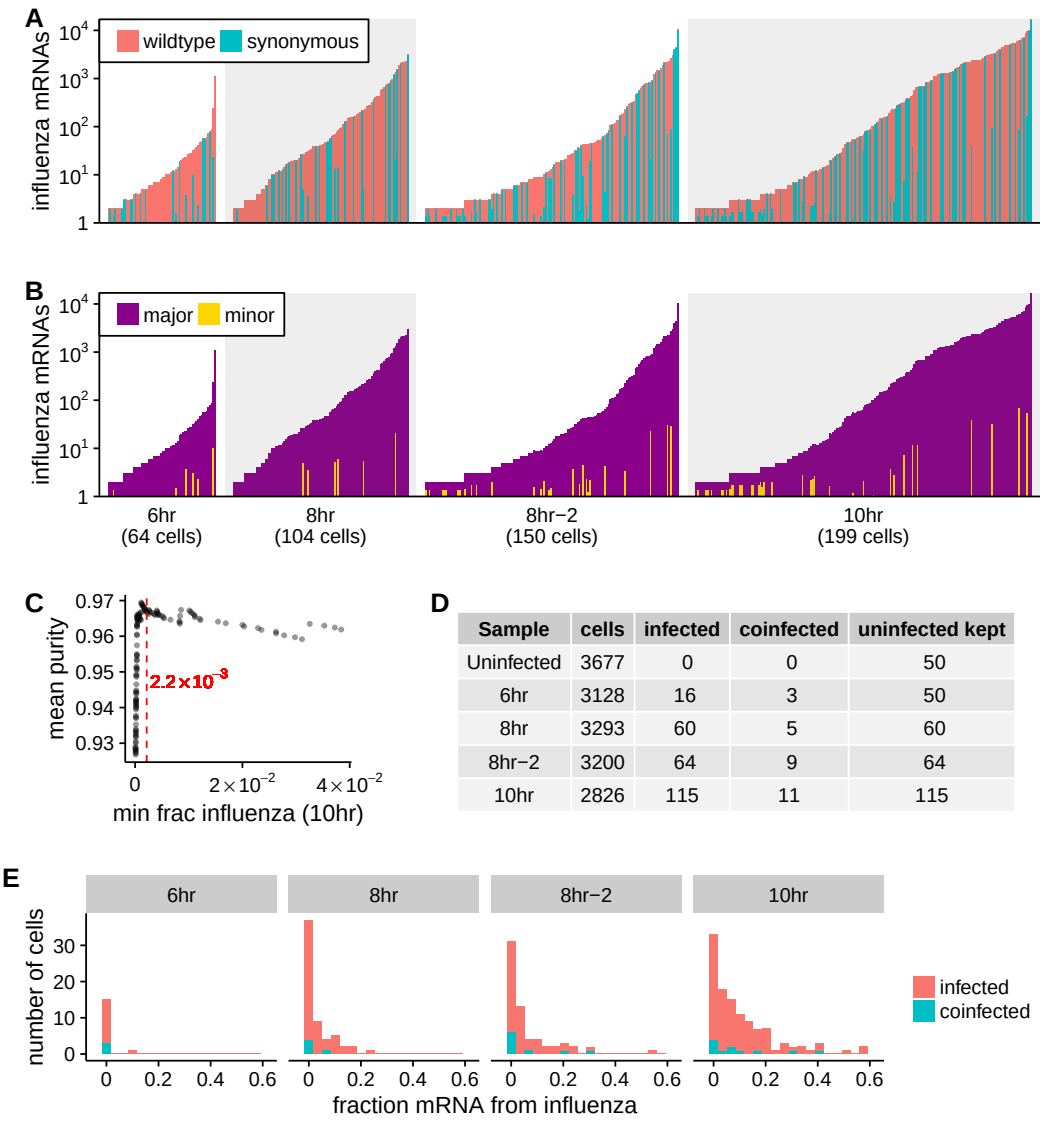
Discussion

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.



Sample	cells	infected	coinfected	uninfected kept
Uninfected	3677	0	0	50
6hr	3128	16	3	50
8hr	3293	60	5	60
8hr−2	3200	64	9	64
10hr	2826	115	11	115

Figure 5. 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.

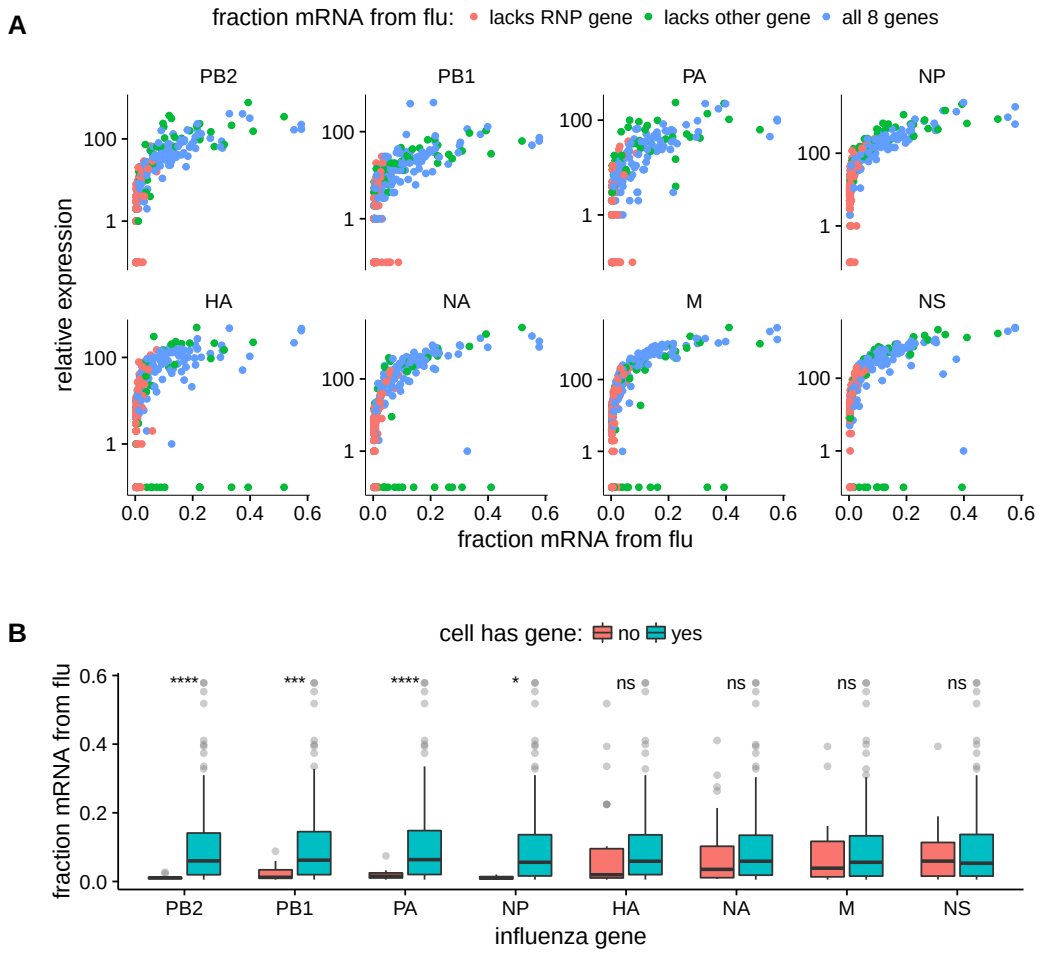


Figure 6. The viral infection burden in individual cells as a function of the amount of each viral gene detected. **(A)** Fraction of mRNAs in each cell derived from virus as a function of the *normalized* expression of each viral gene in that cell. This plot shows that all cells with very high viral 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.

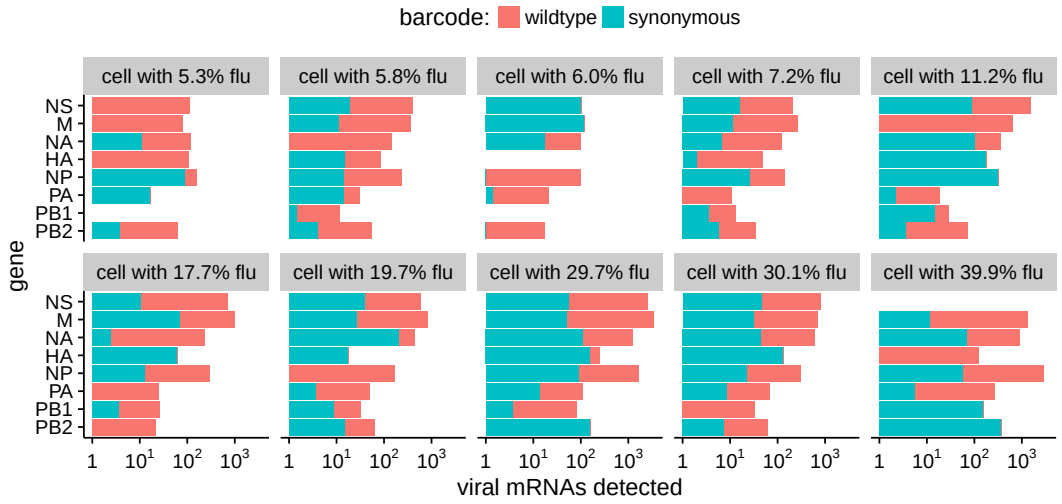


Figure 7. Frequency of each viral gene segment in co-infected cells with at least 5% of their mRNA derived from influenza. The bars indicate the logarithm of the number of each viral mRNA detected, and the bars are colored in proportion to the fraction of those mRNAs that are derived from either wildtype or synonymous barcoded virus.

Figure 7–source data 1. The raw data plotted in this figure are in `p_coinfection.csv`.

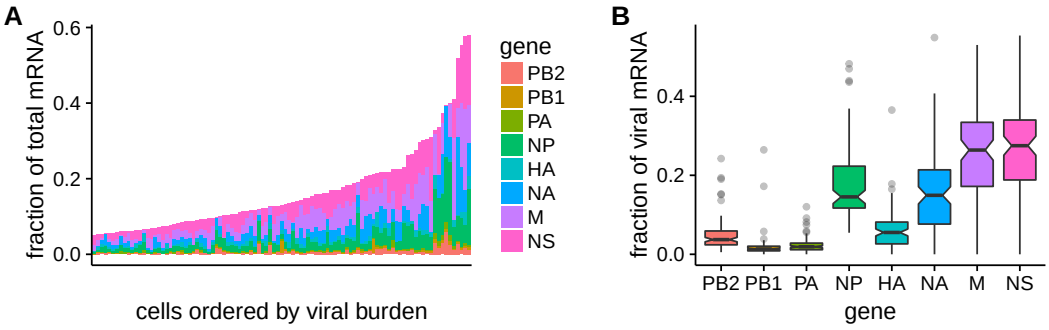


Figure 8. Expression of individual influenza genes in highly infected cells (at least 5% of total mRNA is viral). **(A)** The fraction of total mRNA from each influenza gene for each cell. **(B)** Box plots showing the fraction of viral mRNA per cell that is derived from each influenza gene taken over 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 8–source data 1. The raw data are in `p_flu_expr.csv`.

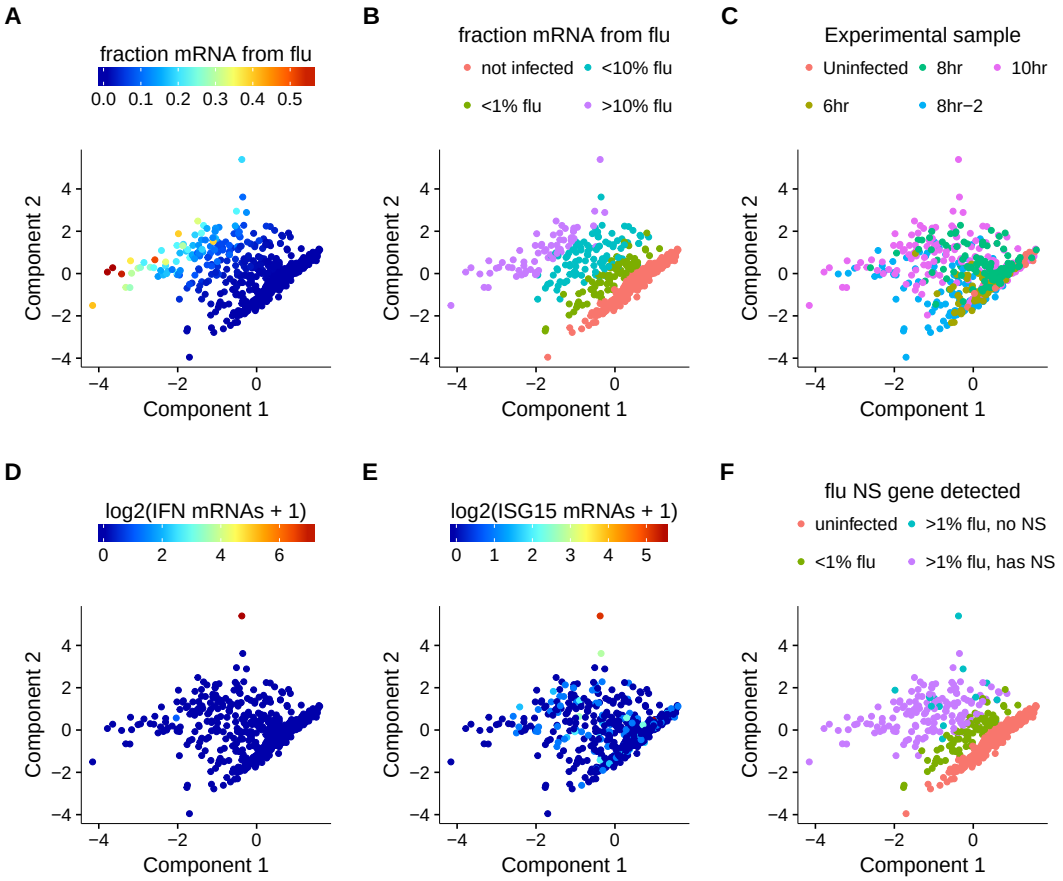


Figure 9. Visual layout of the cells according to "pseudotime". The layout is the same in all panels, but each panel colors the cells according to a different property. **(A), (B)** Cells colored by the fraction of their mRNA that is viral. **(C)** Cells colored by experimental sample. While it is clear 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.

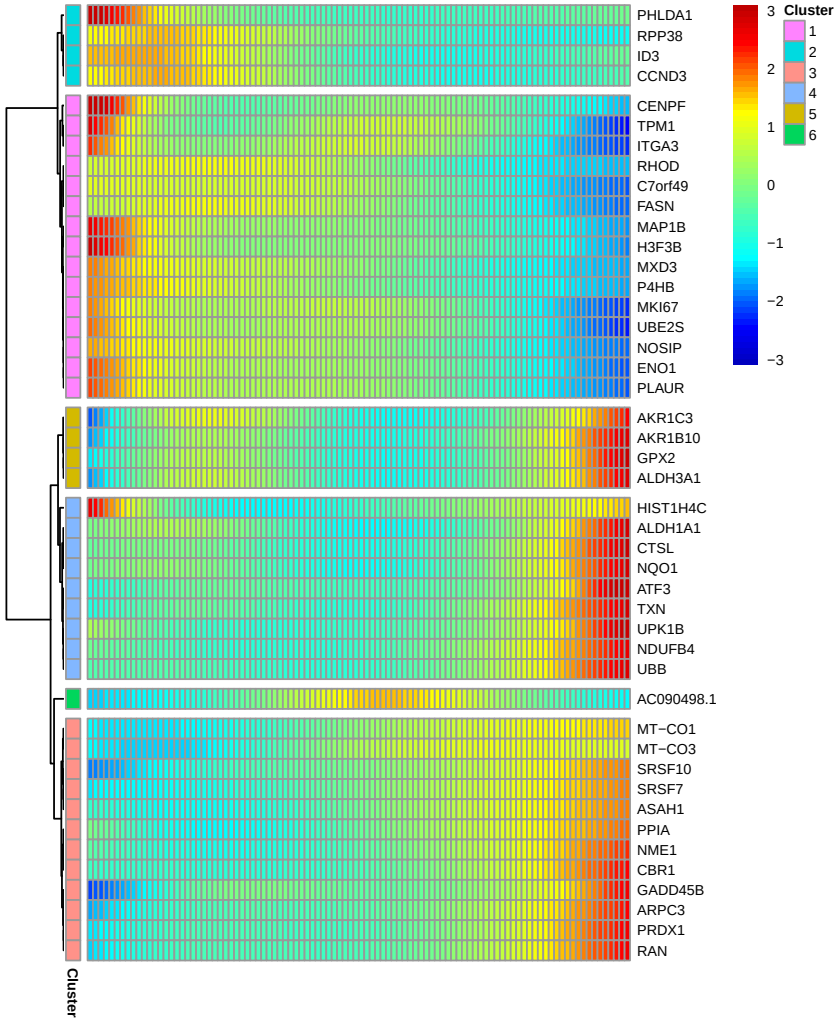


Figure 10. Cellular genes that are differentially expressed with respect to the amount of influenza mRNA in individual cells infected with full influenza virus containing all eight genes. Shown are all genes differentially expressed with $Q < 0.1$.

Figure 10–source data 1. The full results of the differential expression test is in `p_sig_cellular_genes.csv`.

Figure 10–source data 2. The results of a gene-set analysis are in `p_sig_cellular_genes.csv`.

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.

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

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

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

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

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

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

Figure 3–Figure supplement 1. This is a supplementary figure's full caption, which will be used at the end of the manuscript.

Figure 3–Figure supplement 2. This is another supplementary figure.