DEVELOPMENT OF DROPLET MICROFLUDIC TECHNOLOGIES

FOR NEGATIVE-STRAND RNA VIRUSES

by

Mallory Marie Blackwood

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Microbiology and Immunology

MONTANA STATE UNIVERSITY
Bozeman, Montana

May 2025

DEDICATION

This dissertation is dedicated to everyone committed to protecting the integrity of scientific research, evidence-based policy, and critical thinking in our society. Creating a world that is equitable and just, for people and planet, requires empathy and collaboration from us all.

ACKNOWLEDGEMENTS

I would like to acknowledge the support and guidance of my family, friends, and collaborators, without who this work would not be possible. Most importantly I would like to thank my advisor, Dr. Connie Chang, for years of scientific training and personal friendship.

Thank you to my loving partner Hannah for your unwavering kindness, patience, and support. I could not do this without you. Thank you to my parents and my brothers for cheering me along the whole way, I hope to make you proud. Thank you especially to my mother Ann, who taught me it is never too late to start something new. Thank you to my grandparents, Elizabeth and Howard Blackwood, who inspired me to be curious about the world and the people in it. I wish you were here to share this day. Thank you to my friends, old and new, who help me remember that there is nothing more important in life than adventure and laughter. Thank you especially to my friend Patrick who I once promised, and perhaps am contractually obligated, to acknowledge in this document. Thanks for being willing to go down the rabbit hole with me.

I would like to thank all members of the Chang Lab, especially the co-authors of this work. Thank you to Emma Loveday, whose mentorship contributed significantly to this research. Thank you to the Brooke lab at UIUC, the Ke lab at Los Alamos National Laboratory, the Galanis and Sung labs at Mayo Clinic, and the Fields and Keil labs at Montana State. Thank you to the Mayo Clinic Genomics Core and Bioinformatics Core for your contributions to sequencing data acquisition and analysis. Thank you to the members of my committee for your feedback and direction. And thank you to the Montana State Graduate School and Molecular Biosciences Fellowship Program for funding my PhD program.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF TABLES CONTINUED

Table Page

LIST OF TABLES CONTINUED

Table Page

LIST OF FIGURES

Figure Page

LIST OF FIGURES CONTINUED

Figure Page

ABSTRACT

Viruses infect all organisms and play a key role in host population dynamics and evolution. RNA viruses evolve extremely rapidly, infect a wide range of host cell types, and cause severe diseases. The incredible diversity of RNA virus populations, combined with the heterogeneity of host cells they infect, makes studying RNA virus infection quite challenging. Averaging outcomes across RNA virus infection in bulk often misses minor variants and phenotypes that can drive overall population dynamics. Therefore, it is increasingly important to study RNA virus infections at the single-cell level. Droplet microfluidics is a powerful tool for single-cell analysis. By physically isolating individual cells within microfluidic droplets, and manipulating those droplets to conduct various biological assays, we can study RNA virus infection one cell at a time. In this work I describe the establishment of three new droplet microfluidic methods for studying clinically relevant negative-strand RNA viruses, influenza A virus (IAV) and measles virus (MV), at the single-cell level. The first method involves culturing IAV from single infected cells encapsulated in microfluidic droplets. We show that cell viability, genome replication, and viral infectivity in droplets is comparable to bulk infection conditions. The second method builds on this model system by developing a way to directly quantify the number of viral particles produced by each cell. We find that IAV production varies widely between individual cells and across different IAV strains. Together, these platforms provide a high-throughput solution for studying IAV infection at the single-cell level that can be used to explore the effects of antiviral interventions on IAV replication. The third method focuses on viral genome diversity at the single-cell level. We developed a single-cell sequencing approach to study variation in the MV P gene, which contains attenuation markers important for MV function as a cancer immunotherapy treatment. Preliminary data shows the development of custom single-cell sequencing beads to capture the MV P gene attenuation site and generation of Illumina sequencing libraries. Further optimization is required to improve sequencing yield for diversity analysis. Overall, this work introduces novel methodologies for high-throughput single-cell analysis of negative-strand RNA virus infection dynamics and population diversity.

INTRODUCTION

Research Motivation

Negative-strand RNA viruses can be highly pathogenic and cause severe disease in human populations worldwide. RNA genomes exhibit the highest single-nucleotide mutation rate in nature, enabling the frequent emergence of novel virus strains that can evade host immune responses and antiviral treatments. The genetic diversity within negative-strand RNA virus populations also leads to variable infection outcomes at both the cellular and organismal levels. Understanding how single-cell heterogeneity contributes to differences in infection outcome is crucial for developing effective antiviral treatments and vaccines. Droplet microfluidics is a powerful tool for studying single-cell heterogeneity. However, assays must be customized for each new model system. In this study we describe the development of three droplet microfluidic platforms designed to investigate single-cell heterogeneity during infection of influenza A virus (IAV) and measles virus (MV), two clinically relevant negative-strand RNA viruses.

Funding and Collaborations

Research funding was supplied by the Defense Advanced Research Projects Agency (DARPA) grant W911NF-17-2-0034, National Institute of Allergy and Infectious Diseases (NIAID) of the National Institutes of Health (NIH) under award number 1R56AI156137, National Science Foundation (NSF) CAREER grant 1753352, the National Institute of General Medical Sciences (NIGMS) of the National Institutes of Health (NIH) under NOSI NOT-GM-21-031, Montana State IDeA Network of Biomedical Research Excellence (INBRE), and Mayo Clinic internal funding sources. Specific PhD program funding was provided by the Montana

State Graduate School and Molecular Biosciences Fellowship Program. The Brooke lab at UIUC and the Ke lab at Los Alamos National Laboratory collaborated significantly on all influenza projects. The Mayo Clinic Genomics Core and Bioinformatics Core provided data acquisition and analysis support. The Galanis lab and Sung lab at Mayo Clinic collaborated on all measles virus work. The Fields lab and Keil lab at Montana State collaborated on all coronavirus projects. We thank all these funding sources and collaborators for their support.

Dissertation Overview

The remainder of this dissertation is comprised of four chapters and a series of appendices. Chapter 2 reviews current literature on RNA virus diversity and the use of single-cell technologies such as droplet microfluidics to study these diverse populations. We specifically highlight two negative-strand RNA viruses, influenza a virus (IAV) and measles virus (MV), whose relevance in clinical research and biotechnology make them important species to study. Chapter 3 introduces a droplet microfluidic platform for culturing IAV from single-cells. Chapter 4 presents a method for quantifying the number of IAV particles produced from single-cells in drops, expanding the capabilities for IAV single-cell infection study. Chapter 5 outlines a single-cell RNA sequencing platform customized to target a small region of interest on the MV P gene, containing virus attenuation markers, that is relevant to MV function as a cancer immunotherapy. The body of this dissertation is followed by an appendix containing sections A and B. These appendices include (A) supplemental information for the manuscript in chapter 4; (B) supplemental information for the manuscript pre-print in chapter 5.

Chapter 2: "Literature Review"

Chapter Overview: "This dissertation focuses on the development of novel droplet microfluidics techniques for studying highly heterogeneous negative-strand RNA virus populations at the single-cell level. It addresses the need for customized methods to culture viruses in droplets, quantify virus replication within those droplets, and capture the diversity of gene expression profiles that may explain cell-to-cell variation in observed infection dynamics. Here, I will outline the necessity for these types of analysis techniques and review existing methods for studying influenza A virus and measles virus at the single-cell level."

Chapter 3: "Single-cell Infection of Influenza A Virus Using Drop-based Microfluidics"

Chapter Hypothesis: Influenza A virus infection kinetics, such as genome copy number and infectivity titer, remain consistent between single cell infection in microfluidic droplets and bulk infection in traditional tissue culture flasks.

Chapter Abstract: "Drop-based microfluidics has revolutionized single-cell studies and can be applied toward analyzing tens of thousands to millions of single-cells and their products contained within picoliter-sized drops. Drop-based microfluidics can shed insight into single-cell virology, enabling higher-resolution analysis of cellular and viral heterogeneity during viral infection. In this work, individual A549, MDCK, and siat7e cells were infected with influenza A virus (IAV) and encapsulated into 100-μm-size drops. Initial studies of uninfected cells encapsulated in drops demonstrated high cell viability and drop stability. Cell viability of uninfected cells in the drops remained above 75%, and the average drop radii changed by less than 3% following cell encapsulation and incubation over 24 h. Infection parameters were analyzed over 24 h from individually infected cells in drops. The number of IAV viral genomes

and infectious viruses released from A549 and MDCK cells in drops was not significantly different from bulk infection as measured by reverse transcriptase quantitative PCR (RT-qPCR) and plaque assay. The application of drop-based microfluidics in this work expands the capacity to propagate IAV viruses and perform high-throughput analyses of individually infected cells."

Chapter 4: "Influenza A Viral Burst Size from Thousands of Infected Single-cells Using Droplet Quantitative PCR (dqPCR)"

Chapter Hypothesis: Influenza A viral burst size is heterogeneously distributed across individual infected cells within a population, and the population-level average burst size varies between distinct influenza strains.

Chapter Abstract: "An important aspect of how viruses spread and infect is the viral burst size, or the number of new viruses produced by each infected cell. Surprisingly, this value remains poorly characterized for influenza A virus (IAV), commonly known as the flu. In this study, we screened tens of thousands of cells using a microfluidic method called droplet quantitative PCR (dqPCR). The high-throughput capability of dqPCR enabled the measurement of a large population of infected cells producing progeny virus. By measuring the fully assembled and successfully released viruses from these infected cells, we discover that the viral burst sizes for both the seasonal H3N2 and the 2009 pandemic H1N1 strains vary significantly, with H3N2 ranging from 10^1 to 10^4 viruses per cell, and H1N1 ranging from 10^1 to 10^3 viruses per cell. Some infected cells produce average numbers of new viruses, while others generate extensive number of viruses. In fact, we find that only 10% of the single-cell infections are responsible for creating a significant portion of all the viruses. This small fraction produced approximately 60% of new viruses for H3N2 and 40% for H1N1. On average, each infected cell

of the H3N2 flu strain produced 709 new viruses, whereas for H1N1, each infected cell produced 358 viruses. This novel method reveals insights into the flu virus and can lead to improved strategies for managing and preventing the spread of viruses."

Chapter 5: "Droplet Microfluidic Platform for Single-cell Sequencing of Measles Virus During Oncolytic Virotherapy"

Chapter Hypothesis: The measles virus vaccine strain attenuation site, located on the Phosphoprotein gene, can be selectively sequenced from individual infected cells using a custom single-cell RNA sequencing platform.

Chapter Abstract: "As an RNA virus, measles virus (MV) genomes are subject to frequent mutations and selective advantages in diverse host cell environments. During *in vitro* passaging of the MV vaccine strain, two random mutations in the P gene resulted in a new attenuated MV strain with capabilities for cancer immunotherapy. Attenuated MVs can specifically target cancer cells that overexpress the CD46 membrane receptor. MVs infect and kill these cells, simultaneously recruiting a localized immune response to the tumor site. This represents a promising new form of oncology treatment, currently undergoing Phase I and II clinical trials. However, there is considerable heterogeneity in tumor susceptibility to MV infection during these trials. This variability may be related to further mutation of the P gene attenuation site during therapeutic infection. Single-cell analysis of attenuation site diversity could inform the development of MV strains with more consistent infection outcomes and greater success as a cancer immunotherapy. In this work, we develop a customized single-cell RNA sequencing (scRNA-seq) platform to capture the MV P gene attenuation site. Using droplet microfluidics, we synthesize barcoding hydrogel beads (BHBs) with primer sequences specific to

P gene mRNA. These BHBs are then co-encapsulated with single Vero cells infected with attenuated MV. We present preliminary data from two independent replicates that demonstrates the success of our platform in generating Illumina sequencing libraries. This proof-of-concept study will be further developed to optimize protocols for increased sequencing library yield, enabling multiple sequence alignments and diversity analysis of MV P gene expression across thousands of single infected cells."

Chapter 6: "Conclusion"

Chapter Overview: "The results presented in this dissertation expand the use of droplet microfluidics to the study of two clinically relevant negative-strand RNA viruses, IAV and MV. Specifically, we focus on the adaptation of molecular biology assays, such as RT-qPCR and next generation sequencing, to a single-cell platform, for nucleic acid quantification and sequencing from IAV and MV RNA genomes and transcripts. Ultimately, the goal of this work is to develop high-throughput methods that can be used to study various features of RNA virus population diversity and evolution at the single-cell level."

CHAPTER TWO

LITERATURE REVIEW

Overview

This dissertation focuses on the development of novel droplet microfluidics techniques for studying highly heterogeneous negative-strand RNA virus populations at the single-cell level. It addresses the need for customized methods to culture viruses in droplets, quantify virus replication within those droplets, and capture the diversity of gene expression profiles that may explain cell-to-cell variation in observed infection dynamics. Here, I will outline the necessity for these types of analysis techniques and review existing methods for studying influenza A virus and measles virus at the single-cell level.

RNA Virus Population Diversity

Viruses are taxonomically classified by features of their genome structure, including nucleic acid type (RNA or DNA), sense (positive or negative), and strand (single stranded or double stranded) [1]. These genome features determine the molecular mechanisms used by a virus to complete the central dogma of biology, converting genetic information from nucleic acid to messenger RNA to protein [2]. Viruses can be further classified by genome packaging (linear or circular, single or multiple segments), capsid symmetry (icosahedral, helical, or random), envelope (presence / absence), genome and amino acid sequence similarity, order of genes on the genome, mechanism of mRNA synthesis, and enzyme usage [3].

One of the major classification groups are RNA viruses, which are characterized by populations with high genetic diversity and rapid evolution. Genetic diversity in RNA virus populations is mediated by frequent mutation of genomes across large population sizes [4],[5]. RNA genomes have the highest single nucleotide mutation rate in nature (Fig 1) due to the low-fidelity RNA dependent RNA polymerase required for replication [6]-[9]. This is evolutionarily advantageous because it increases the probability of generating a genome variant with enhanced fitness against novel selection pressures [10],[11] in heterogeneous host cell environments [12]. As such, RNA viruses exist as populations of closely related, but genetically distinct, viral genome variants referred to as viral quasispecies [13]. The high genetic diversity of viral quasispecies enables evasion of host immune responses and antiviral treatments, as well as infection across a broad host range [14]-[25].

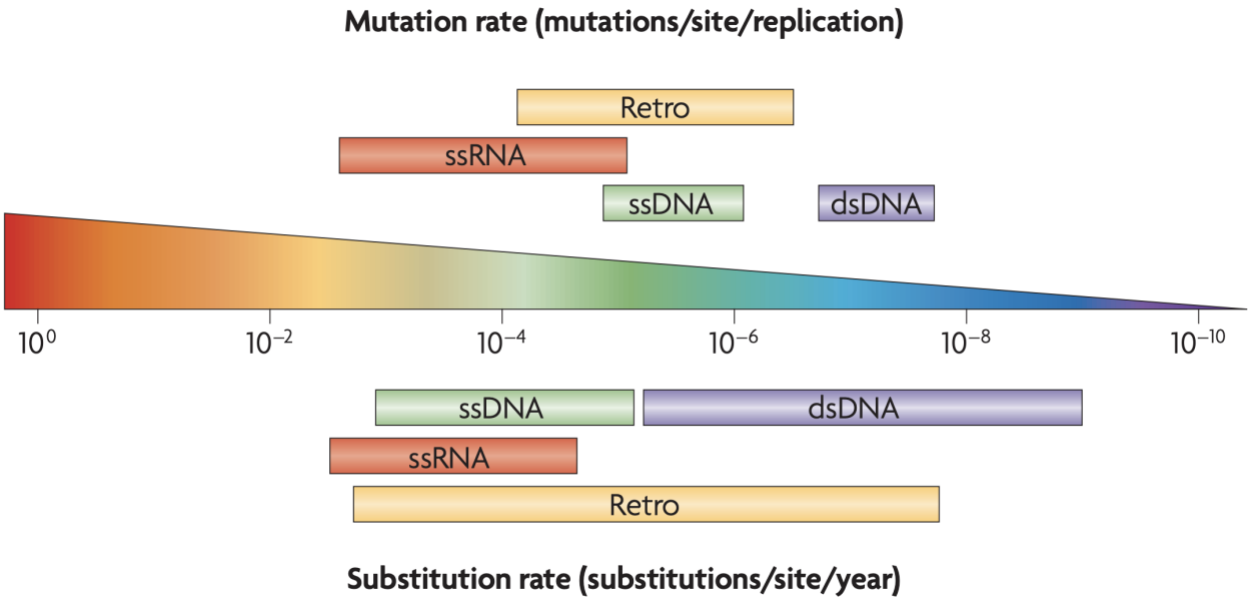


Figure 1. Comparison between viral mutation and substitution rates. *Adapted from [6], with permission from the journal.* The ranges of mutation rates, given as mutations per site per round of replication, for viruses with different genomic architectures are summarized in the upper part of the figure. The ranges of average substitution rates, given as substitutions per site per year, are shown in the lower part of the figure.

Many RNA viruses have specifically impacted human populations through repeated disease outbreaks. Two such RNA viruses were studied in this work: influenza A virus (IAV) and measles virus (MV). Both viruses have single-stranded negative-sense RNA genomes, a group that consists of other widespread and deadly viruses such as Ebola virus, Rabies virus, and Lassa fever virus. Negative-sense RNA genomes are packaged into a helical nucleocapsid that is immediately transcribed into positive-sense messenger RNA (mRNA), by a RNA dependent RNA polymerase (RdRP) enzyme packaged within the virion, upon infection. After the genome is transcribed into mRNA, viral protein translation and genome replication can begin. This strategy for RNA genome replication introduces single point mutations that contribute to high levels of genetic diversity in the population.

While IAV and MV have the same genome type, their infection cycles and disease presentations vary significantly. Both viruses spread from host to host through aerosolization of respiratory droplets. However, the rate of transmission varies dramatically. Viral transmission potential is measured by reproductive number (R_0), which quantifies the average number of secondary infections produced by one infected individual in a fully susceptible population. R_0 essentially describes how easily a disease could spread through an unprotected population. Low R_0 values are associated with low rates of spread, and high R_0 values are associated with high rates of spread. Interestingly, IAVs have a lower R_0 value (median 1.28-1.80) while MVs have a higher R_0 value (median 12-18), that may vary widely by region [26].

There are also key differences in the recommended vaccination regimens to protect against IAV and MV infection. Individuals should be vaccinated for IAV annually to protect against seasonal emergence of new antigenic variants, to which we do not have pre-existing

immunity. However, because IAV has a low rate of person-to-person transmission, fewer individuals need to be vaccinated to achieve heard immunity. In contrast, there is only a single antigenic variant of MV in circulation. Meaning that once a person is vaccinated, they will be protected with life-long immunity against MV infection. However, high rates of MV transmission from person-to-person do mean that a large proportion of the population needs to be vaccinated against MV to achieve heard immunity. These key differences between IAV and MV infection dynamics highlights an important nuance in negative-strand RNA virus research – the relationship between genetic diversity and antigenic variation. Specific to this work, both IAV and MV populations contain high genetic diversity due to RdRp induced point mutations that cause genetic drift over time. However, IAV populations typically have high antigenic diversity while MV populations have low antigenic diversity [6]. This is because IAV populations are subject to larger genetic shifts in the population that create antigenic variation. The reasons for this are rooted in the different mechanisms for genetic diversity generation in each virus species, which will be further discussed below.

Clinically Relevant Negative-Strand RNA Viruses

Influenza A Virus (IAV)

IAV is a negative-strand RNA virus belonging to the family *Orthomyxoviridae*. It has an enveloped virion that ranges from 80-120 nm in size and carries a linear single-stranded negative-sense RNA genome that is 10-15kb in length [3]. IAV genomes exhibit an extremely high level of genetic diversity [27]. The genome is split into 8 segments that each encode one or two viral proteins. Frequent genome point mutation and reassortment between co-infecting

viruses results in rapid evolution and continued emergence of novel strains, posing a significant threat to public health through seasonal disease epidemics and global pandemic events [3].

There are four primary subtypes of influenza virus (A, B, C, and D) that circulate within multiple animal species and can move between animal species. IAV is the most common subtype found in human populations [28]. IAV infects epithelial cells in the upper respiratory tract and causes respiratory disease with symptoms including fever, sore throat, cough, headache, and muscle pain. IAVs can be further classified by the amino acid sequence of two surface glycoproteins, Hemagglutinin (HA) and Neuraminidase (NA), which facilitate virion entry into host cells through interactions with sialic acid residues [3]. It is also the unique structure of different the HA and NA proteins that is used to classify IAV into different species subtypes, or strains. The primary IAV strains currently circulating in human populations are H1N1 and H3N2 [29]. The H1N1 strain originated from a spill over event from pigs to humans in 2009, and the H3N2 strain from a similar spillover event from birds to pigs to humans in 1998 [30]. Novel IAV strains such as H1N1 and H3N2 have been the source of multiple global pandemics, leading to 50 million human deaths, over the last century. In addition to historic pandemic events, IAV remains a significant public health threat by causing seasonal epidemics that average 25 million cases, 20,000 deaths, and 10 billion dollars in medical costs each year in the United States [31].

The repeated emergence of novel IAV strains is a result of extreme genetic diversity which allows for rapid adaptation to selective pressures that directly impact transmissibility, pathogenicity, and virulence. Although there are vaccine and antiviral treatments against IAVs, new strains emerge each year that can escape established host immune response. Development of

more effective vaccines and antiviral treatments to reduce transmission and pathogenesis of IAVs will require a better understanding of the mechanisms that drive their diversity.

IAVs exist as a genetically diverse quasispecies with significant heterogeneity in viral gene expression [32],[33] and host innate immune activation [34] that varies both within and between host cell types and IAV strains [35]. This diversity is attributed to unique features of the IAV genome. Like other RNA viruses, IAV genomes are characterized by high mutational frequency during each round of replication (averaging 10^{-5} mutations per nucleotide per replication cycle) [6],[9]. In addition to point mutation, IAV genomes also acquire large deletion regions during replication that produce defective viral genomes (DVGs) [36], [37]. DVGs are missing large regions of protein encoding sequence, while maintaining the sequences required for replication and packaging [38]-[40]. The truncated length of DVGs means they are selectively replicated and packaged over full-length WT segments [41],[42] and often outcompete WT virus [43]-[45] to accumulate in the population. The presence of DVGs means that 90-99% of IAV virions will contain incomplete genomes [46] and require co-infection with wild type (WT) particles to replicate [37]. During co-infection, mutated genome segments derived from multiple viruses are frequently packaged into a single progeny virion [47]. This process is called genome reassortment and underlies large genetic shifts in the IAV population that accelerate adaptive evolution [48]. As a result, IAVs respond quickly to changing host environments through repeated zoonotic emergence of novel strains [30],[49], evasion of established host immunity [47], and resistance to antiviral drugs and vaccine treatments [50].

Measles Virus (MV)

Measles viruses (MV) is a negative-strand RNA virus belonging to the family *Paramyxoviridae*. MV has an enveloped virion with a diameter of 150-300 nm that contains a non-segmented genome of linear single-stranded negative-sense RNA measuring 15-18 kb in length [3]. MV RNA genomes have a high single nucleotide mutation rate, averaging 10^{-5} mutations per nucleotide per replication cycle *in vitro* [6],[51] and 10^{-3} mutations per nucleotide per year during genomic surveillance [52]. However, unlike other RNA viruses, MV populations do not undergo significant genetic and antigenic shifts. There is only one wild type MV serotype which circulates across a narrow host range and remains antigenically stable over time [52].

Symptoms of MV infection, which mostly occurs in children, include a characteristic red blotchy rash that is accompanied by fever, red eyes, cough, runny nose, and spots in the mouth. The attachment and entry of MV into host cells is mediated by two envelope glycoproteins, hemagglutinin (H) and fusion (F) proteins. Wild Type MV infects immune cells such as B lymphocytes, T lymphocytes, dendritic cells, and monocytes via the CD150 'SLAM' cell surface protein [3]. Infection of immune cells may contribute to the high pathogenicity and transmission associated with MV infection. The MV vaccine strain, however, has been adapted to bind to many different cell types via the CD46 cell surface protein [3]. Thanks to widespread vaccination, MV has not been a public health concern in the United States for most of the 21st century but has posed a significant threat across Africa and Asia [54]. Development of the Edmonston live-attenuated vaccine in 1961 decreased case numbers from 4 million per year to 500 per year in the United States. Continued genomic surveillance of MV shows limited antigenic variation or emerging strains in the population over time [55]. As a result, serial

passaging derivatives of the original 1961 vaccine strain have remained highly effective at conferring lifelong immunity to vaccinated individuals.

The MV genome contains 6 genes that are translated to produce 1 to 3 proteins each, for a total of 8 proteins encoded in the genomes [3]. This work focuses on measurement of the MV phosphoprotein (P) gene, whose product is a component of the MV RNA polymerase complex [3]. The P gene also encodes two non-structural C and V proteins. Together, the P, C, and V proteins help regulate the host immune response to infection by blocking interferon signaling [56]-[57]. Specifically, the P protein tyrosine 110 is required for interferon blocking [56]. Mutations within the P gene, acquired by serial passaging of vaccine strain derivatives, reduce MVs ability to block host interferon responses [58]. This type of mutation further attenuates the virus so that it cannot cause disease but can still elicit a strong immune response. This, along with an affinity for the CD46 receptor [59], which is over expressed in cancerous tumor tissue [60]-[62], has made MV useful as an exciting new form of oncolytic virotherapy to treat cancerous tumors [63]-[66].

Immunotherapy is widely used in cancer treatment, where specific antigens are introduced to train a patient's own immune system to recognize and eliminate cancerous cells and tumors. Many types of immunotherapies exist, including checkpoint inhibitors, adoptive cell therapy, monoclonal antibodies, immune system modulators, cancer vaccines, and oncolytic virotherapies. Oncolytic virotherapy works by direct injection into cancerous tumors, where viruses selectively replicate in and kill cancer cells through syncytia formation and lysis, while also recruiting a localized immune response to the tumor site [66]. MV is highly advantageous as an oncolytic virotherapy because of the existence of a safe and effective vaccine strain, that has a

high affinity for CD46 receptors that are overexpressed on cancerous tumor tissue and can be further attenuated by mutations in the P gene to allow for host interferon signaling to proceed during infection and recruit a localized immune response to tumor sites.

Single-cell Virology

Characterizing the infection dynamics of RNA viruses, like IAV and MV, is challenging due to extreme genomic and phenotypic heterogeneity in their populations. Most research on RNA viruses is conducted using bulk assays, which average outcomes across a population. While informative, bulk studies do not capture important subsets of heterogenous RNA virus populations. This challenge can be overcome by observing viral infection at the single-cell level [67].

Many methods exist for single-cell isolation, the most common being flow cytometry, laser microdissection, manual cell picking, limiting dilution, and microfluidics [68]. Microfluidics is perhaps the most widely used method of single-cell analysis, utilizing microwell, valve, and drop based technologies (Fig 2) [67]. Once single-cells are isolated, they can be analyzed using various imaging [69] and multi-omics [70] techniques. These techniques have been applied to single-cell virology to address questions about progeny production, virus kinetics, defective viral genomes, and viral gene expression [71]-[72]. Investigation of these infection dynamics at the single-cell level has revealed extreme heterogeneity in response to antiviral agents [73], which is critical for the development of highly effective treatments.

Previous investigations of single-cell IAV infections have found high variability in viral gene expression [32]-[33] (Heldt 2015), host innate immune activation [34], defective viral

genome content [40] and viral burst size [74]. Single-cell IAV infection dynamics have also shown to be MOI dependent [75] and vary between host cell types and IAV strains [35].

Single-cell studies of MV infection are more limited and have primarily focused on the portion of the viral lifecycle where MV infection moved from epithelial cells to immune cells. These studies have helped identify which immune cells are most susceptible to MV infection [76], the cellular functions that enable movement of infection centers from respiratory epithelium to immune cells [77]-[78], and the role of interferon signaling in stopping the dissociation of infection centers [79].

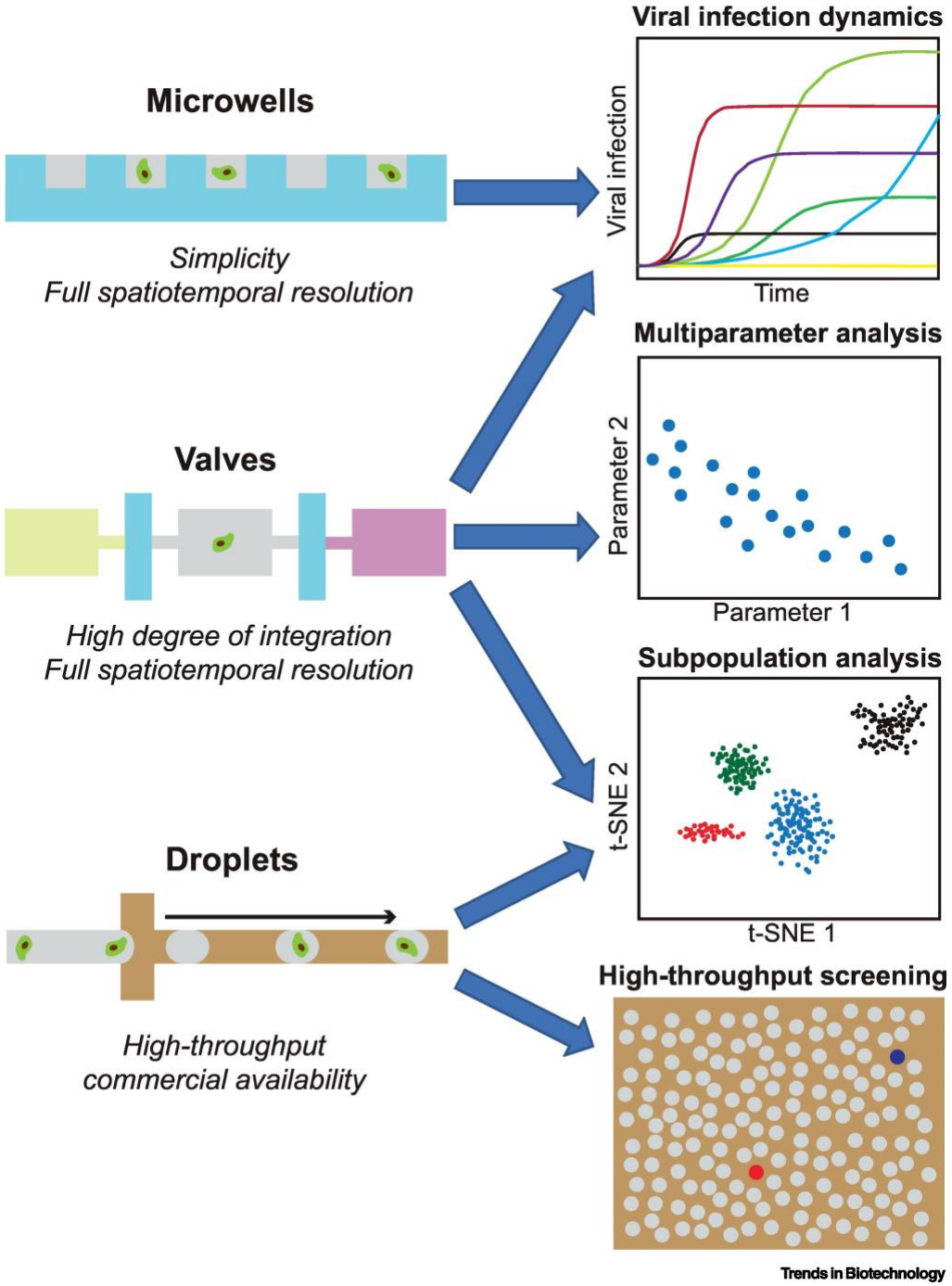


Figure 2. Single-Cell Microfluidic Technologies and Their Typical Applications in Virology. *Adapted from [67], with permission from the journal.* Microwell-, valve-, and droplet-based strategies are applied in virology studies. Features and typical application opportunities of these methods are presented. Abbreviations: t-SNE 1 and t-SNE 2, dimension 1 and dimension 2 of the data space with t-distributed stochastic neighbor embedding (t-SNE) technique for dimensionality reduction.

Droplet Microfluidics

One powerful method of single-cell analysis is droplet microfluidics, which physically isolates single-cells into droplets ranging 10-200 μm in diameter. Droplets can be generated at a rate of thousands per second and modified downstream to provide miniaturized and high-throughput replacements of many laboratory assays [80] addressing multi omics questions (Fig 3) [81]. Droplet microfluidics has been widely applied to the field of single-cell virology [67], to track infection dynamics, identify subpopulations with important genotype or phenotype, and perform high-throughput screening of both. It is a critical tool for clinical research, serving as infectious disease diagnostics [82] for detection of pathogens like RNA viruses [83] such as IAV and MV [84]. Additionally, droplet microfluidics helps to uncover more basic properties about the diversity of viral replication [85], infectivity [86], genome recombination [87],[88], and evolutionary trajectory [89]-[91]. In this section we will discuss the basics of droplet microfluidics and describe its application to three important laboratory assays (cell culture, RT-qPCR, and next generation sequencing) which are prevalent in the remainder of this work.

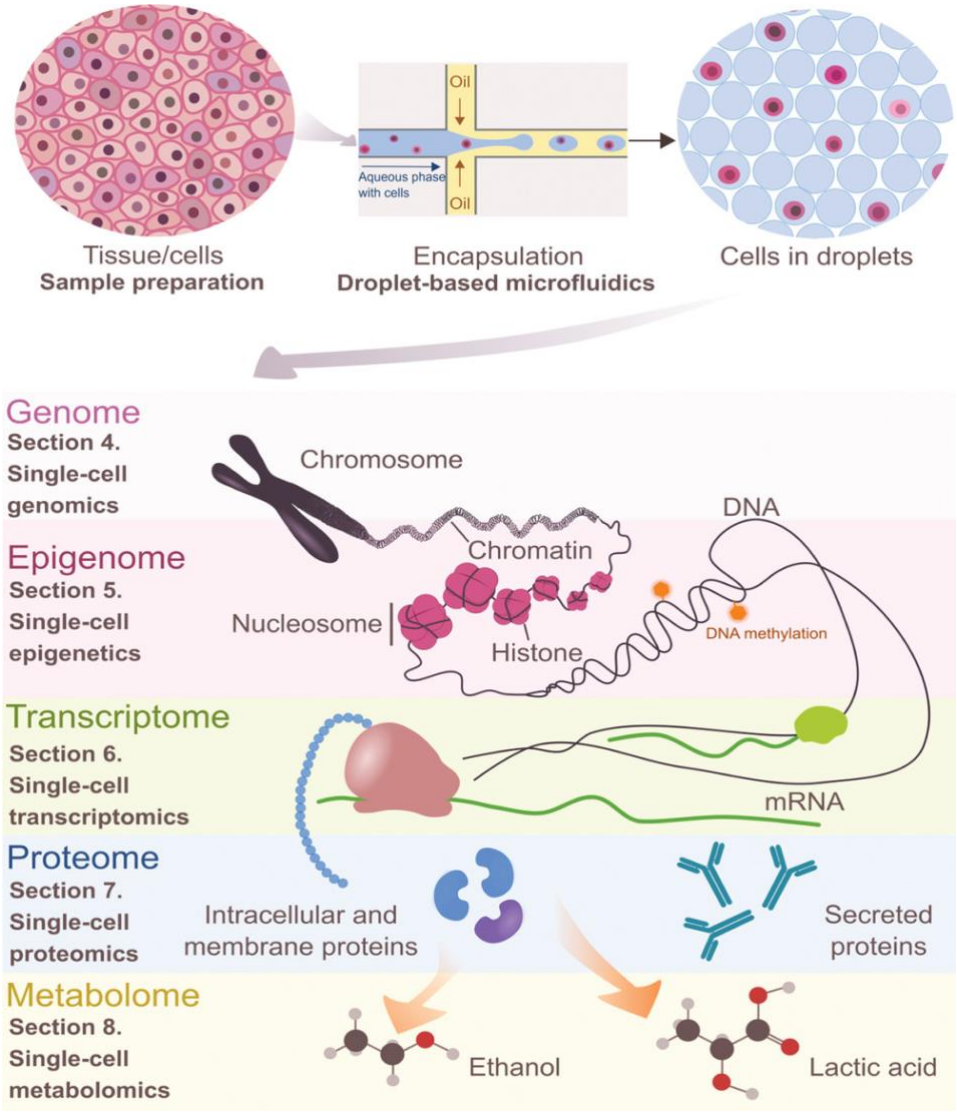


Figure 3. Single-cell Analysis Using Droplet Microfluidics. *Adapted from [81], with permission from the journal.* Overview of droplet microfluidics approaches to gain information on all key processes in individual cells. Technologies are categorized in sections as indicated in the figure.

Droplet Generation

Microfluidic droplets at their most basic form are water-in-oil emulsions. The aqueous phase and the oil phase do not mix within the droplet or between different droplets. Enabling full isolation of encapsulated contents from neighboring droplets. Droplets are formed by flowing small liquid volumes into microfluidic devices at high-speed using either a flow or pressure-

based systems. Microfluidic devices are typically made of polydimethylsiloxane (PDMS) polymer chips that contain micron scale channels and junctions molded from SU8 features printed onto silicon wafers [92]. Fabricated PMDS chips are then bound to a glass slide using plasma surface treatment, and the channels are made hydrophobic or hydrophilic, depending on the fluids being used. This is the basic device fabrication protocol for the devices described in this work. However, many methods and materials exist for microfluidic device fabrication.

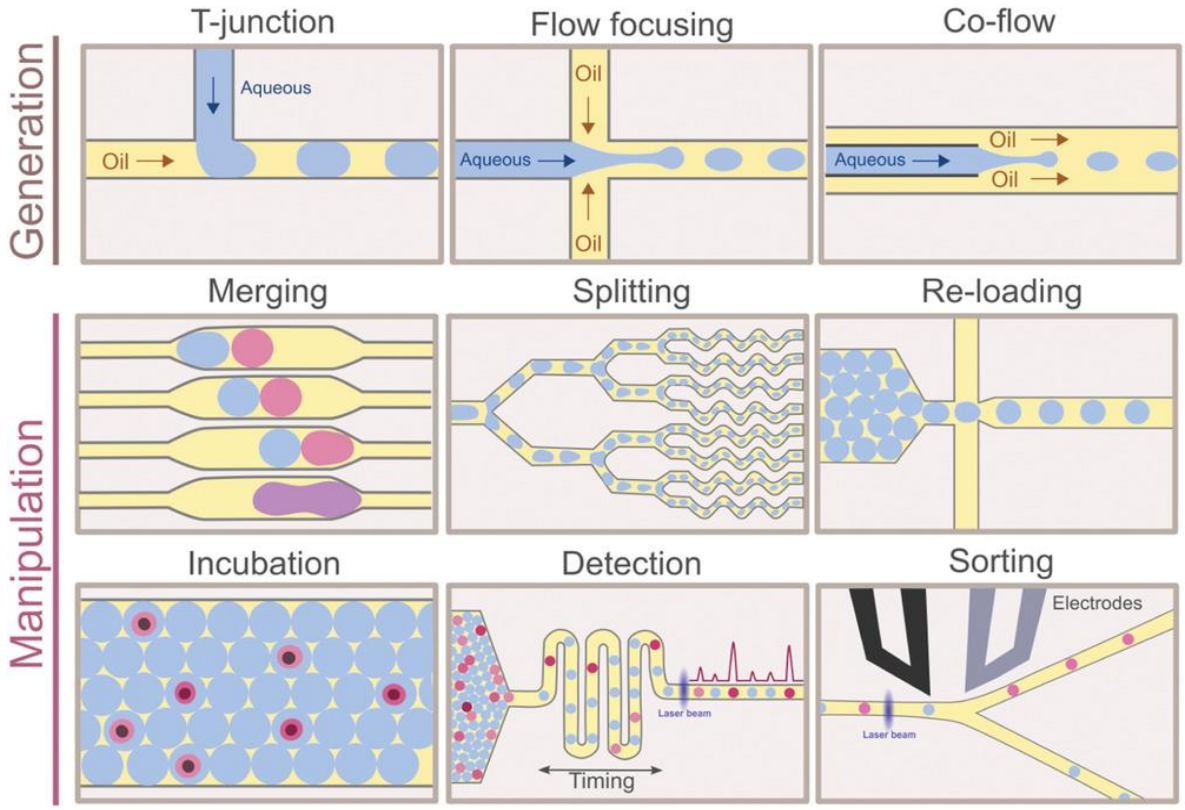


Figure 4. Microfluidic drop maker geometries and manipulation modules applied in single-cell workflows. *Adapted from [81], with permission from the journal.* Droplets can be generated using microfluidic devices using three different geometries: T-junction, flow focusing, and co-flow. After droplet generation, microfluidic modules are available for droplet manipulation: merging of droplets with different contents, splitting of bigger droplets into droplets with smaller volume, re-loading of the emulsion into a new device, static on-chip incubation of the emulsion at different temperatures, detection of (fluorescence) signal in droplets, and sorting droplets of interest (using e.g., electric field).

Custom device design enables complex fluid flow and manipulation, allowing for the miniaturization of almost any bulk assay onto a microfluidic chip. Devices can be designed to include multilayered channels or embedded electric elements [93]. Droplets can be split, merged, reinjected, heated, detected, and sorted (Fig 4) [81]. Droplet microfluidics has revolutionized single-cell biology by providing a platform that decreases the cost and increases the throughput of traditional laboratory assays. Single-cells can be encapsulated into drops at 10 kHz rates, which allows for millions of drops to be produced in a short period of time [94]. This yields a corresponding increase in statistical power of the conclusions made from droplet microfluidic assays. Three droplet microfluidic assays will be discussed in detail below – virus culture, droplet quantitative PCR, and single-cell sequencing – that are critical to the research described in this dissertation.

Virus Culture in Droplets

The most fundamental method to virology research is virus culture. To culture viruses from single-cells encapsulated into microfluidic droplets we must ensure that cells remain viable and virus replication and infection is comparable to bulk infection conditions [85],[86],[95]. Without these basic quality control measures, single-cell assays cannot be compared to bulk assay counterparts. One important parameter is the biocompatibility of oil used for droplet generation. We use fluorinated oils and surfactants for single-cell encapsulation. These fluorinated oils are nontoxic to cells and allow for gas exchange between the aqueous phase and the atmosphere [96]. Another important factor is media nutrient concentration within the droplets. Small volumes used during single-cell culture in drops may require optimization of media components, compared to media used during bulk cell culture. Additionally, virus culture

typically requires that infected cells are incubated at warm temperatures (averaging at 35°C). This presents a challenge to droplet stability, where drops may coalesce above room temperature. This condition can be overcome by again optimizing the contents of single-cell infection media and may require the addition of chemicals meant to stabilize the drop interface [97]. However, these chemical additions may also disrupt cell viability and virus infection cycles, therefore testing of concentration gradients is typically required.

Droplet Quantitative PCR

One of the most widely used methods for quantifying species abundance is measurement of nucleic acid concentration through quantitative polymerase-chain reaction (qPCR) assay. To measure the concentration of RNA viruses in a sample, we can target regions of the RNA genome for detection through qPCR [98],[99]. When utilized in a droplet microfluidic system, qPCR assays increase in speed, sensitivity, and throughput. This technique is commercially available through digital droplet PCR (ddPCR), which splits a bulk sample into drops and counts the number of drops that produce a positive fluorescence read-out through qPCR [100]. The number of positive drops corresponds to target concentration in the starting sample.

We can also use qPCR in drops to measure viral RNA concentration from single infected cells [74]. Previous work on RT-qPCR of single infected cells (limiting dilution) shows that IAV progeny virus output and intracellular viral RNA concentration can vary up to 3 orders of magnitude [33],[101]. However, these single-cell methods are limited in through-put. This dissertation describes an improved upon system for single-cell RT-qPCR, which samples viral RNA directly from single-cells in drops. Our method, termed droplet quantitative PCR (dqPCR) [74], splits secreted extracellular virus from the host cell and merges it with qPCR mix. Drops

are thermocycled off -chip and the fluorescence intensity of individual drops are read in series using high-speed flow-based detection on a microfluidic chip [93]. The fluorescence intensity of each drop is converted to a pre-split RNA copy number using a sigmoidal curve fitting model. The model relies on generation of an RNA standard droplet amplification curve library that is specific to a target sequence of interest, to relate droplet fluorescence at a given cycle number back to a starting RNA concentration [97]. The dqPCR method enables quantification of viral RNA from tens of thousands of single-cell infections. The goal of dqPCR is to quantify heterogeneity in viral RNA replication across single-cell infections and compare results to existing bulk infection data.

Single-cell Sequencing

Much of the work in the field of single-cell virology has focused on single-cell RNA sequencing (scRNA-seq) technologies [102]. scRNA-seq is typically used to quantify diversity in gene expression at a single-cell level and identifies important subpopulations that are often masked by bulk RNA sequencing methods. Understanding the true heterogeneity of RNA expression during viral infection may inform fundamental concepts of viral spread and persistence, activation of immune responses, and the design of antiviral therapeutics.

Sc-RNAseq relies on microfluidic technologies [103]. Most scRNA-seq platforms are designed for transcriptomic sequencing of cellular messenger RNA (mRNA). The technology was originally developed by the inDrop [104] and Dropseq [105] protocols, which were later commercialized by 10x Genomics. These platforms utilize microfluidic devices to construct barcoded hydrogel beads (BHBs) carrying DNA primers with sequences for mRNA capture and barcoding, then to co-encapsulate BHBs with single-cells in droplets. Once isolated in

microfluidic drops, reverse transcription of cellular mRNA yields cDNA barcoded by drop of origin. This barcoded cDNA (BC-cDNA) is processed for Illumina sequencing, producing tens of thousands of scRNA-seq libraries derived from individual cells.

Since the development of Dropseq and InDrop, and subsequent commercialization by 10x Genomics, many other scRNA-seq methods have been commercialized. These new technologies allow for scRNA-seq of genomes, epigenomes, and proteomes, in addition to traditional transcriptomic sequencing. NanoString has expanded the field of single-cell transcriptomics to spatial monitoring of gene expression in tissue layers. Mission Bio specializes in single-cell genomics and multiomics, using bead-based microfluidic technologies to simultaneously capture DNA and proteins. Scale Biosciences and Parse Biosciences are microfluidic-free scRNA-seq technologies that build sequencing barcodes directly onto cells themselves, reducing the technical skill, equipment need, and cost of the assay. Even the industry leader in next generation sequencing, Illumina, has recently released their own protocol for multiomics scRNA-seq. Continued research is improving upon each step of the scRNA-seq procedure, like faster sequencing bead construction [106] and microfluidics-free encapsulation of single-cells and sequencing beads into droplets [107]. The wide array of scRNA-seq platforms currently available and under development ensures this technique will remain a staple in biomedical research for decades to come.

References Cited

1. Baltimore, D. Viral genetic systems. *Trans. N. Y. Acad. Sci.* **33**, 327–332 (1971).

2. Crick, F. Central dogma of molecular biology. *Nature* **227**, 561–563 (1970).

3. Acheson, N. Fundamentals of molecular virology 2nd edition - NLM Catalog - NCBI. https://www.ncbi.nlm.nih.gov/nlmcatalog/101551557 (2011).

4. Dolan, P. T., Whitfield, Z. J. & Andino, R. Mapping the Evolutionary Potential of RNA Viruses. *Cell Host Microbe* **23**, 435–446 (2018).

5. Dolan, P. T., Whitfield, Z. J. & Andino, R. Mechanisms and concepts in RNA virus population dynamics and evolution. *Annu. Rev. Virol.* **5**, 69–92 (2018).

6. Duffy, S., Shackelton, L. A. & Holmes, E. C. Rates of evolutionary change in viruses: patterns and determinants. *Nat. Rev. Genet.* **9**, 267–276 (2008).

7. Sanjuán, R., Nebot, M. R., Chirico, N., Mansky, L. M. & Belshaw, R. Viral Mutation Rates. *J. Virol.* **84**, 9733–9748 (2010).

8. Sanjuán, R. From molecular genetics to phylodynamics: evolutionary relevance of mutation rates across viruses. *PLoS Pathog.* **8**, e1002685 (2012).

9. Drake, J. W. Rates of spontaneous mutation among RNA viruses. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 4171–4175 (1993).

10. Dutta, R. N., Rouzine, I. M., Smith, S. D., Wilke, C. O. & Novella, I. S. Rapid adaptive amplification of preexisting variation in an RNA virus. *J. Virol.* **82**, 4354–4362 (2008).

11. Hayden, E. J., Ferrada, E. & Wagner, A. Cryptic genetic variation promotes rapid evolutionary adaptation in an RNA enzyme. *Nature* **474**, 92–95 (2011).

12. Jones, J. E., Le Sage, V. & Lakdawala, S. S. Viral and host heterogeneity and their effects on the viral life cycle. *Nat. Rev. Microbiol.* **19**, 272–282 (2021).

13. Eigen, M. Viral quasispecies. *Scientific American* **269**, 42–49 (1993).

14. Holland, J. *et al.* Rapid evolution of RNA genomes. *Science* **215**, 1577–1585 (1982).

15. Domingo, E. *RNA Genetics: Volume III: Variability of RNA Genomes*. (CRC Press, London, England, 2020). doi:10.1201/9781351076449.

16. Holland, J. J., De La Torre, J. C. & Steinhauer, D. A. RNA virus populations as quasispecies. *Curr. Top. Microbiol. Immunol.* **176**, 1–20 (1992).

17. Domingo, E. *et al.* Basic concepts in RNA virus evolution. *FASEB J.* **10**, 859–864 (1996).

18. Domingo, E. & Holland, J. J. RNA virus mutations and fitness for survival. *Annu. Rev. Microbiol.* **51**, 151–178 (1997).

19. Domingo, E. Viruses at the edge of adaptation. *Virology* **270**, 251–253 (2000).

20. Vignuzzi, M., Stone, J. K., Arnold, J. J., Cameron, C. E. & Andino, R. Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. *Nature* **439**, 344–348 (2006).

21. Lauring, A. S. & Andino, R. Quasispecies theory and the behavior of RNA viruses. *PLoS Pathog.* **6**, e1001005 (2010).

22. Domingo, E., Sheldon, J. & Perales, C. Viral quasispecies evolution. *Microbiol. Mol. Biol. Rev.* **76**, 159–216 (2012).

23. Andino, R. & Domingo, E. Viral quasispecies. *Virology* **479–480**, 46–51 (2015).

24. Poirier, E. Z. & Vignuzzi, M. Virus population dynamics during infection. *Curr. Opin. Virol.* **23**, 82–87 (2017).

25. Sanjuán, R. The Social Life of Viruses. *Annu Rev Virol* (2021) doi:10.1146/annurev-virology-091919-071712.

26. Rodpothong, P. & Auewarakul, P. Viral evolution and transmission effectiveness. *World J. Virol.* **1**, 131–134 (2012).

27. Brooke, C. B. Population Diversity and Collective Interactions during Influenza Virus Infection. *J. Virol.* **91**, (2017).

28. Ghedin, E. *et al.* Large-scale sequencing of human influenza reveals the dynamic nature of viral genome evolution. *Nature* **437**, 1162–1166 (2005).

29. Finkelman, B. S. *et al.* Global patterns in seasonal activity of influenza A/H3N2, A/H1N1, and B from 1997 to 2005: viral coexistence and latitudinal gradients. *PLoS One* **2**, e1296 (2007).

30. Mostafa, A., Abdelwhab, E. M., Mettenleiter, T. C. & Pleschka, S. Zoonotic Potential of Influenza A Viruses: A Comprehensive Overview. *Viruses* **10**, (2018).

31. Molinari, N.-A. M. *et al.* The annual impact of seasonal influenza in the US: measuring disease burden and costs. *Vaccine* **25**, 5086–5096 (2007).
32. Russell, A. B., Trapnell, C. & Bloom, J. D. Extreme heterogeneity of influenza virus infection in single cells. *Elife* **7**, (2018).

33. Heldt, F. S., Kupke, S. Y., Dorl, S., Reichl, U. & Frensing, T. Single-cell analysis and stochastic modelling unveil large cell-to-cell variability in influenza A virus infection. *Nat. Commun.* **6**, 8938 (2015).

34. Russell, A. B., Elshina, E., Kowalsky, J. R., te Velthuis, A. J. W. & Bloom, J. D. Single-Cell Virus Sequencing of Influenza Infections That Trigger Innate Immunity. *J. Virol.* **93**, (2019).

35. Sun, J. *et al.* Single cell heterogeneity in influenza A virus gene expression shapes the innate antiviral response to infection. *PLoS Pathog.* **16**, e1008671 (2020).

36. Von Magnus, P. Incomplete Forms of Influenza Virus. in *Advances in Virus Research* (eds. Smith, K. M. & Lauffer, M. A.) vol. 2 59–79 (Academic Press, 1954).

37. Jacobs, N. T. *et al.* Incomplete influenza A virus genomes occur frequently but are readily complemented during localized viral spread. *Nat. Commun.* **10**, 3526 (2019).

38. Davis, A. R., Hiti, A. L. & Nayak, D. P. Influenza defective interfering viral RNA is formed by internal deletion of genomic RNA. *Proc. Natl. Acad. Sci. U. S. A.* **77**, 215–219 (1980).

39. Vignuzzi, M. & López, C. B. Defective viral genomes are key drivers of the virus–host interaction. *Nature microbiology* (2019).

40. Wang, C. *et al.* Cell-to-Cell Variation in Defective Virus Expression and Effects on Host Responses during Influenza Virus Infection. *MBio* **11**, (2020).

41. Duhaut, S. D. & McCauley, J. W. Defective RNAs inhibit the assembly of influenza virus genome segments in a segment-specific manner. *Virology* **216**, 326–337 (1996).

42. 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. *J. Virol.* **71**, 2138–2145 (1997).

43. Huang, A. S. & Baltimore, D. Defective viral particles and viral disease processes. *Nature* **226**, 325–327 (1970).

44. Nayak, D. P., Tobita, K., Janda, J. M., Davis, A. R. & De, B. K. Homologous interference mediated by defective interfering influenza virus derived from a temperature-sensitive mutant of influenza virus. *J. Virol.* **28**, 375–386 (1978).

45. Akpinar, F., Timm, A. & Yin, J. High-Throughput Single-Cell Kinetics of Virus Infections in the Presence of Defective Interfering Particles. *J. Virol.* **90**, 1599–1612 (2016).

46. Brooke, C. B. *et al.* Most influenza a virions fail to express at least one essential viral protein. *J. Virol.* **87**, 3155–3162 (2013).

47. Scholtissek, C. Molecular evolution of influenza viruses. *Virus Genes* **11**, 209–215 (1995).

48. Raghwani, J., Thompson, R. N. & Koelle, K. Selection on non-antigenic gene segments of seasonal influenza A virus and its impact on adaptive evolution. *Virus Evol* **3**, vex034 (2017).

49. Taubenberger, J. K. & Kash, J. C. Influenza virus evolution, host adaptation, and pandemic formation. *Cell Host Microbe* **7**, 440–451 (2010).

50. Corder, B. N., Bullard, B. L., Poland, G. A. & Weaver, E. A. A Decade in Review: A Systematic Review of Universal Influenza Vaccines in Clinical Trials during the 2010 Decade. *Viruses* **12**, (2020).

51. Schrag, S. J., Rota, P. A. & Bellini, W. J. Spontaneous mutation rate of measles virus: direct estimation based on mutations conferring monoclonal antibody resistance. *J. Virol.* **73**, 51–54 (1999).

52. Kühne, M., Brown, D. W. G. & Jin, L. Genetic variability of measles virus in acute and persistent infections. *Infect. Genet. Evol.* **6**, 269–276 (2006).

53. Bellini, W. J. & Rota, P. A. Genetic diversity of wild-type measles viruses: implications for global measles elimination programs. *Emerg. Infect. Dis.* **4**, 29–35 (1998).

54. Moss, W. J. & Griffin, D. E. Global measles elimination. *Nat. Rev. Microbiol.* **4**, 900–908 (2006).

55. Kasibhatla, S. M., Vaidya, S. R., Kale, M. M. & Kulkarni-Kale, U. Phylogenomics and evolution of measles virus. in *Phylogenomics* 391–413 (Elsevier, 2024).

56. Devaux, P., von Messling, V., Songsungthong, W., Springfeld, C. & Cattaneo, R. Tyrosine 110 in the measles virus phosphoprotein is required to block STAT1 phosphorylation. *Virology* **360**, 72–83 (2007).

57. Devaux, P., Priniski, L. & Cattaneo, R. The measles virus phosphoprotein interacts with the linker domain of STAT1. *Virology* **444**, 250–256 (2013).

58. Ohno, S., Ono, N., Takeda, M., Takeuchi, K. & Yanagi, Y. Dissection of measles virus V protein in relation to its ability to block alpha/beta interferon signal transduction. *J. Gen. Virol.* **85**, 2991–2999 (2004).

59. Condack, C., Grivel, J.-C., Devaux, P., Margolis, L. & Cattaneo, R. Measles virus vaccine attenuation: suboptimal infection of lymphatic tissue and tropism alteration. *J. Infect. Dis.* **196**, 541–549 (2007).

60. Anderson, B. D., Nakamura, T., Russell, S. J. & Peng, K.-W. High CD46 receptor density determines preferential killing of tumor cells by oncolytic measles virus. *Cancer Res.* **64**, 4919–4926 (2004).

61. Ong, H. T. *et al.* Oncolytic measles virus targets high CD46 expression on multiple myeloma cells. *Exp. Hematol.* **34**, 713–720 (2006).

62. Russell, S. J. & Peng, K. W. Measles virus for cancer therapy. *Curr. Top. Microbiol. Immunol.* **330**, 213–241 (2009).

63. Aref, S., Bailey, K. & Fielding, A. Measles to the Rescue: A Review of Oncolytic Measles Virus. *Viruses* **8**, 294 (2016).

64. Msaouel, P. *et al.* Clinical trials with oncolytic measles virus: Current status and future prospects. *Curr. Cancer Drug Targets* **18**, 177–187 (2018).

65. Mühlebach, M. D. Measles virus in cancer therapy. *Curr. Opin. Virol.* **41**, 85–97 (2020).

66. Engeland, C. E. & Ungerechts, G. Measles virus as an oncolytic immunotherapy. *Cancers (Basel)* **13**, 544 (2021).

67. Liu, W., He, H. & Zheng, S.-Y. Microfluidics in single-cell virology: Technologies and applications. *Trends Biotechnol.* **38**, 1360–1372 (2020).

68. Gross, A. *et al.* Technologies for Single-Cell Isolation. *Int. J. Mol. Sci.* **16**, 16897–16919 (2015).

69. Mattiazzi Usaj, M., Yeung, C. H. L., Friesen, H., Boone, C. & Andrews, B. J. Single-cell image analysis to explore cell-to-cell heterogeneity in isogenic populations. *Cell Syst* **12**, 608–621 (2021).

70. Lee, J., Hyeon, D. Y. & Hwang, D. Single-cell multiomics: technologies and data analysis methods. *Exp. Mol. Med.* **52**, 1428–1442 (2020).

71. Adams, S. K., Ducharme, G. E. & Loveday, E. K. All the single cells: if you like it then you should put some virus on it. *J. Virol.* **98**, e0127323 (2024).

72. Bost, P. & Drayman, N. Dissecting viral infections, one cell at a time, by single-cell technologies. *Microbes Infect.* **26**, 105268 (2024).

73. Guo, F. *et al.* Single-Cell Virology: On-Chip Investigation of Viral Infection Dynamics. *Cell Rep.* **21**, 1692–1704 (2017).

74. Thomas, M. M. *et al.* Influenza A viral burst size from thousands of infected single cells using droplet quantitative PCR (dqPCR). *PLoS Pathog.* **20**, e1012257 (2024).

75. Ramos, I. *et al.* Innate Immune Response to Influenza Virus at Single-Cell Resolution in Human Epithelial Cells Revealed Paracrine Induction of Interferon Lambda 1. *J. Virol.* **93**, (2019).

76. Acklin, J. A. *et al.* Immunological landscape of human lymphoid explants during measles virus infection. *JCI Insight* **9**, e172261 (2024).

77. Hippee, C. E. *et al.* Measles virus exits human airway epithelia within dislodged metabolically active infectious centers. *PLoS Pathog.* **17**, e1009458 (2021).

78. Hippee, C. E. *et al.* Epithelial-to-mesenchymal transition and live cell extrusion contribute to measles virus release from human airway epithelia. *J. Virol.* **99**, e0122024 (2025).

79. Durnell, L. A. *et al.* Interferon-independent processes constrain measles virus cell-to-cell spread in primary human airway epithelial cells. *Microbiol. Spectr.* e0136123 (2023).

80. Guo, M. T., Rotem, A., Heyman, J. A. & Weitz, D. A. Droplet microfluidics for high-throughput biological assays. *Lab Chip* **12**, 2146–2155 (2012).

81. Matuła, K., Rivello, F. & Huck, W. T. S. Single-Cell Analysis Using Droplet Microfluidics. *Adv Biosyst* **4**, e1900188 (2020).

82. Lehnert, T. & Gijs, M. A. M. Microfluidic systems for infectious disease diagnostics. *Lab Chip* **24**, 1441–1493 (2024).

83. Zhu, H., Fohlerová, Z., Pekárek, J., Basova, E. & Neužil, P. Recent advances in lab-on-a-chip technologies for viral diagnosis. *Biosens. Bioelectron.* **153**, 112041 (2020).

84. Krokhine, S., Torabi, H., Doostmohammadi, A. & Rezai, P. Conventional and microfluidic methods for airborne virus isolation and detection. *Colloids Surf. B Biointerfaces* **206**, 111962 (2021).

85. Fischer, A. E. *et al.* A high-throughput drop microfluidic system for virus culture and analysis. *J. Virol. Methods* **213**, 111–117 (2015).

86. Tao, Y. *et al.* Rapid, targeted and culture-free viral infectivity assay in drop-based microfluidics. *Lab Chip* **15**, 3934–3940 (2015).

87. Zhang, H. *et al.* Isolation and Analysis of Rare Norovirus Recombinants from Coinfected Mice Using Drop-Based Microfluidics. *J. Virol.* **89**, 7722–7734 (2015).

88. Tao, Y. *et al.* Artifact-Free Quantification and Sequencing of Rare Recombinant Viruses by Using Drop-Based Microfluidics. *Chembiochem* **16**, 2167–2171 (2015).

89. Agresti, J. J. *et al.* Ultrahigh-throughput screening in drop-based microfluidics for directed evolution. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 4004–4009 (2010).

90. Rotem, A. *et al.* Evolution on the Biophysical Fitness Landscape of an RNA Virus. *Mol. Biol. Evol.* **35**, 2390–2400 (2018).

91. Wang, Y. *et al.* Directed evolution: Methodologies and applications. *Chem. Rev.* **121**, 12384–12444 (2021).

92. Duffy, D. C., McDonald, J. C., Schueller, O. J. & Whitesides, G. M. Rapid prototyping of microfluidic systems in poly(dimethylsiloxane). *Anal. Chem.* **70**, 4974–4984 (1998).

93. Mazutis, L. *et al.* Single-cell analysis and sorting using droplet-based microfluidics. *Nat. Protoc.* **8**, 870–891 (2013).

94. Theberge, A. B. *et al.* Microdroplets in microfluidics: an evolving platform for discoveries in chemistry and biology. *Angew. Chem. Int. Ed Engl.* **49**, 5846–5868 (2010).

95. Loveday, E. K., Sanchez, H. S., Thomas, M. M. & Chang, C. B. Single-Cell Infection of Influenza A Virus Using Drop-Based Microfluidics. *Microbiol Spectr* e0099322 (2022).

96. Holtze, C. *et al.* Biocompatible surfactants for water-in-fluorocarbon emulsions. *Lab Chip* **8**, 1632–1639 (2008).

97. Loveday, E. K., Zath, G. K., Bikos, D. A., Jay, Z. J. & Chang, C. B. Screening of Additive Formulations Enables Off-Chip Drop Reverse Transcription Quantitative Polymerase Chain Reaction of Single Influenza A Virus Genomes. *Anal. Chem.* **93**, 4365–4373 (2021).

98. Mackay, I. M., Arden, K. E. & Nitsche, A. Real-time PCR in virology. *Nucleic Acids Res.* **30**, 1292–1305 (2002).

99. Clementi, M. *et al.* Quantitative PCR and RT-PCR in virology. *PCR methods and applications* **2**, 191–191 (1993).

100. Hindson, B. J. *et al.* High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. *Anal. Chem.* **83**, 8604–8610 (2011).

101. Kupke, S. Y. *et al.* Single-Cell Analysis Uncovers a Vast Diversity in Intracellular Viral Defective Interfering RNA Content Affecting the Large Cell-to-Cell Heterogeneity in Influenza A Virus Replication. *Viruses* **12**, (2020).

102. Ciuffi, A., Rato, S. & Telenti, A. Single-Cell Genomics for Virology. *Viruses* **8**, (2016).

103. Zhou, W.-M. *et al.* Microfluidics applications for high-throughput single cell sequencing. *J. Nanobiotechnology* **19**, 312 (2021).

104. Klein, A. M. *et al.* Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* **161**, 1187–1201 (2015).

105. Macosko, E. Z. *et al.* Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015).

106. Delley, C. L. & Abate, A. R. Modular barcode beads for microfluidic single cell genomics. *Sci. Rep.* **11**, 10857 (2021).

107. Clark, I. C. *et al.* Microfluidics-free single-cell genomics with templated emulsification. *Nat. Biotechnol.* **41**, 1557–1566 (2023).

CHAPTER THREE

SINGLE-CELL INFECTION OF INFLUENZA A VIRUS USING DROP-BASED MICROFLUIDICS

Contribution of Authors and Co-Authors

Manuscript in Chapter 3

Author: Emma K. Loveday

Contributions: Designed studies, Performed Experiments, Analyzed data, Wrote Manuscript.

Co-Author: Humberto S. Sanchez

Contributions: Performed Experiments, Analyzed data, Wrote Manuscript.

Co-Author: Mallory M. Thomas

Contributions: Performed Experiments, Analyzed data, Wrote Manuscript.

Co-Author: Connie B. Chang

Contributions: PI on the Study, Acquired funding, Designed studies, Wrote Manuscript.

Manuscript Information

Emma K. Loveday, Humberto S. Sanchez, Mallory M. Thomas, Connie B. Chang

Microbiology Spectrum

Status of Manuscript:
☐ Prepared for submission to a peer-reviewed journal
☐ Officially submitted to a peer-reviewed journal
☐ Accepted by a peer-reviewed journal
☒ Published in a peer-reviewed journal

American Society for Microbiology

Volume 10, Number 5, October 26th, 2022

DOI 10.1128/spectrum.00993-22

AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Spectrum

RESEARCH ARTICLE

Single-Cell Infection of Influenza A Virus Using Drop-Based Microfluidics

Emma Kate Loveday,[a,b] Humberto S. Sanchez,[a,b] Mallory M. Thomas,[a,c] Connie B. Chang[a,b,d]

[a]Center for Biofilm Engineering, Montana State University, Bozeman, Montana, USA
[b]Department of Chemical and Biological Engineering, Montana State University, Bozeman, Montana, USA
[c]Microbiology and Cell Biology, Montana State University, Bozeman, Montana, USA
[d]Department of Physiology and Biomedical Engineering, Mayo Clinic, Rochester, Minnesota, USA

ABSTRACT Drop-based microfluidics has revolutionized single-cell studies and can be applied toward analyzing tens of thousands to millions of single cells and their products contained within picoliter-sized drops. Drop-based microfluidics can shed insight into single-cell virology, enabling higher-resolution analysis of cellular and viral heterogeneity during viral infection. In this work, individual A549, MDCK, and siat7e cells were infected with influenza A virus (IAV) and encapsulated into 100-μm-size drops. Initial studies of uninfected cells encapsulated in drops demonstrated high cell viability and drop stability. Cell viability of uninfected cells in the drops remained above 75%, and the average drop radii changed by less than 3% following cell encapsulation and incubation over 24 h. Infection parameters were analyzed over 24 h from individually infected cells in drops. The number of IAV viral genomes and infectious viruses released from A549 and MDCK cells in drops was not significantly different from bulk infection as measured by reverse transcriptase quantitative PCR (RT-qPCR) and plaque assay. The application of drop-based microfluidics in this work expands the capacity to propagate IAV viruses and perform high-throughput analyses of individually infected cells.

IMPORTANCE Drop-based microfluidics is a cutting-edge tool in single-cell research. Here, we used drop-based microfluidics to encapsulate thousands of individual cells infected with influenza A virus within picoliter-sized drops. Drop stability, cell loading, and cell viability were quantified from three different cell lines that support influenza A virus propagation. Similar levels of viral progeny as determined by RT-qPCR and plaque assay were observed from encapsulated cells in drops compared to bulk culture. This approach enables the ability to propagate influenza A virus from encapsulated cells, allowing for future high-throughput analysis of single host cell interactions in isolated microenvironments over the course of the viral life cycle.

KEYWORDS microfluidics, influenza, single cell, drops

Single-cell studies of viral infection enable high-resolution examination of heterogeneous virus populations. Differential selective pressures in both the host cell and virus population lead to variability in virus replication and production that enables antiviral escape, zoonotic spillover events, and changes in virulence or pathogenesis (1). Influenza A virus (IAV) is a negative-strand RNA virus with populations containing high genetic diversity due to its segmented genome, rapid replication rate, and low-fidelity RNA-dependent RNA polymerase (RdRp) (1). As such, IAV infection results in a diverse swarm of unique variants exhibiting heterogeneous genotypes and phenotypes (2, 3).

Cutting-edge technologies such as drop-based microfluidics (4–8) and single-cell sequencing (9–14) are enabling higher-resolution analysis of the effects of cellular heterogeneity on viral infection dynamics. Single-cell virology studies have revealed a more detailed snapshot of the heterogeneous kinetics of virus production in relation to innate immune activation and

Editor Abimbola O. Kolawole, Wright State University

Address correspondence to Connie B. Chang, connie.chang@montana.edu.

The authors declare no conflict of interest.

Received 13 April 2022
Accepted 22 August 2022

responses (15–18) and the variability of viral gene expression between individually infected cells (9–12). Heterogeneous IAV populations can be examined using methods that capture diversity at the single-cell level. Such methods include microfluidic techniques to perform single-cell transcriptomics (9–11) or isolation of cells using limiting dilutions in well plates to assess single infection events (12, 13, 16, 19) or by using two-dimensional (2D) microfluidic chambers (15). While these studies have identified large variations in total viral mRNA expressed, infectious virus produced, and host transcriptional response to infection, the low-throughput methods were limited to analyzing hundreds to a few thousand individual cells (12, 13, 15, 20). To process tens of thousands or up to millions of single cells, a high-throughput continuous flow method such as fluorescence-activated cell sorting (FACS) can be utilized to study virus infection (9–11, 14). Yet FACS sorting is mostly performed at early time points postinfection to ensure that virus spread is reduced and that the timing of infection, viral gene expression, and cellular response to infection are comparable from cell to cell (9–11, 14, 17). In contrast, analysis of infected cells from bulk cultures at later time points postinfection can be difficult due to virus spread.

A method to study single cells at high throughput is drop-based microfluidics, which offers the ability to compartmentalize and rapidly assay individual cells (6, 21–25). A microfluidic device is used to create microscale aqueous drops that are surrounded by an immiscible oil and stabilized with a biocompatible surfactant (26). Compared to larger-scale *in vitro* culturing methods where cells are grown in flasks or well plates, drop-based microfluidics creates millions of discrete bioreactors that house single cells in which viruses can replicate (4, 7), enabling viral replication events to be analyzed independently. Furthermore, these compartmentalized cells in drops allow for single-cell analysis at any time point postinfection since virus spread cannot occur between neighboring drops. Thus, drops provide a way to analyze virus replication, production, and cellular responses of single cells at multiple time points.

To date, drop-based viral infections have only been performed with murine norovirus (MNV-1), a positive-sense RNA virus in a small icosahedral capsid that lacks an outer envelope and is extremely stable at a range of environmental conditions, using cell lines adapted to spinner cultures or already grown in suspension (4, 7, 8, 27). Drop-based microfluidics was used to study infectivity (4, 7) and recombination (8) of MNV-1, with data demonstrating the ability to successfully encapsulate and infect mammalian cells within drops (4, 6–8). Expanding the applicability of drop-based microfluidics to other viruses, specifically for enveloped viruses such as IAV, and the ability to encapsulate a broad range of host cells would greatly increase capacity for single-cell virology studies.

In this work, drop-based microfluidics was applied toward culturing and studying IAV infections at the single-cell level. Three different cell lines, alveolar basal epithelial (A549) cells, Madin-Darby Canine Kidney (MDCK) cells, and MDCK plus human *siat7e* gene (siat7e) cells (28, 29), were tested for their viability and ability to support IAV infection in drops compared to bulk tissue culture. Drop stability, cell loading, and viability were tested within 100-μm-diameter drops and quantified after 24 h of incubation, sufficient time for one productive round of IAV infection. Additionally, this study demonstrates the use of drop-based microfluidics to culture and propagate enveloped viruses, which generally have reduced environmental stability. The cells were infected with A/California/07/2009 (H1N1) at a multiplicity of infection (MOI) of 0.1 to compare virus production over multiple time points between infected cells on standard tissue culture plates and encapsulated in drops. Our findings demonstrate that standard adherent cell lines can be used in drops to propagate IAV, thereby expanding single-cell capabilities for studying viral infections, which, until recently, has only been demonstrated for nonadherent, spinner-adapted host cells (4, 7, 8).

The drop-based microfluidic methods developed in this work will enable future studies of single-cell IAV infections using multiple cell lines. The methods expand the application of drop-based microfluidics from the nonenveloped positive-sense MNV-1 (4, 7, 8) to include the enveloped and negative-sense IAV, thus broadening the scope of single-cell virology assays. Additionally, this work can serve as a blueprint for testing, comparing, and implementing drop-based microfluidics methods in future single-cell studies. With the recent interest and development of single-cell omic technologies for

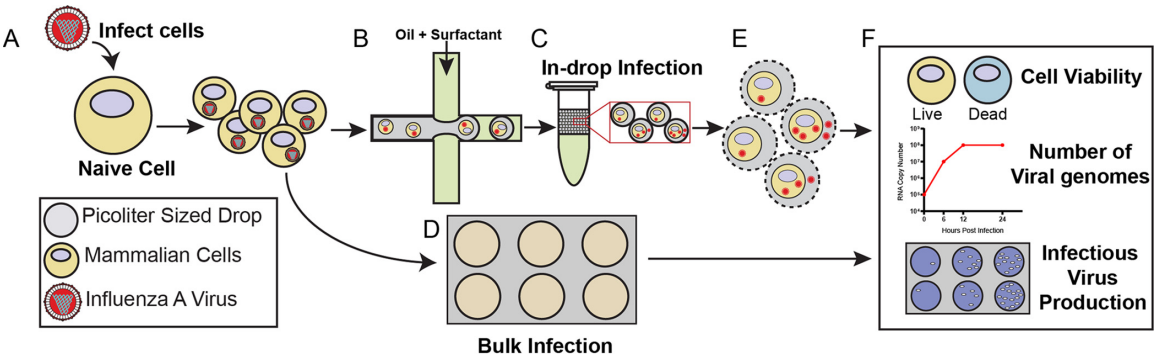


FIG 1 Graphical workflow. (A) Cells were infected with IAV. (B) A suspension of infected cells was encapsulated into drops using a fluorinated oil continuous phase. (C) Drops were incubated for 24 h at 37°C and 5% CO_2. (D) A portion of infected cells was replated onto standard tissue culture dishes to recapitulate a bulk infection and incubated for 24 h at 37°C and 5% CO_2. (E) To analyze virus production, the drops were broken and pooled, and cells and/or viral supernatant were recovered. (F) Cells and virus from drop and bulk infections were analyzed using LIVE/DEAD staining to determine cell viability, RT-qPCR to determine viral genome copy number, and plaque assays to determine viral infectivity.

studying viral infections (13–15, 20, 30–36), we expect this work to further expand capabilities and increase applications of single-cell technologies in virology.

RESULTS

Single-cell virus infection using drop-based microfluidics. A schematic of the virus infection workflow comparing drops to bulk is outlined in Fig. 1. Naive cells were infected with IAV at an MOI of 0.1 prior to encapsulation such that most cells were infected with a single infectious virus particle (Fig. 1A). Following inoculation, cells were dissociated from the tissue culture plates and processed for encapsulation into 100-μm drops (Fig. 1B), while bulk samples were replated (Fig. 1D). Both encapsulated and bulk cells were similarly processed to ensure that trypsinization from the tissue culture plates and subsequent centrifugation and washing did not interfere with infection kinetics. A 0 h bulk and drop sample was obtained immediately following replating or encapsulation. The remaining bulk and drop samples were incubated for 24 h at 37°C to allow for a minimum of one round of virus replication (Fig. 1C and D). At 24 hours post infection (hpi), viral supernatant was collected from both the bulk and drop samples for analysis of viral genome copies and infectious progeny via reverse transcriptase quantitative PCR (RT-qPCR) and plaque assay. To sample viral supernatant from encapsulated infections, the drops were placed in the freezer and, after thawing, treated with 20% 1H,1H,2H,2H-perfluoro-1-octanol (PFO) in HFE7500 to break the emulsions and release the cells and viral supernatant for collection (Fig. 1E). To assess cell viability, drops were not frozen and were broken with 20% PFO in HFE7500 only. We then compared cell viability and number of viral genomes and infectious progeny from different cell lines maintained in bulk culture and encapsulated in drops (Fig. 1F).

Validating IAV infection in different cell lines. Before performing IAV infection in drops, IAV infection was studied in bulk using three different cell lines, A549, MDCK, and siat7e cells. Both the A549 and MDCK cells are anchorage dependent and are commonly used to propagate IAV and investigate viral infection phenotypes. The siat7e cells are a modified MDCK cell line that expresses the *siat7e* gene (ST6GalNac V), which is important for cellular adhesion, allowing the cells to grow in a suspension culture, and they are distinctive from the standard MDCK-Siat1 cells used commonly for IAV production (28, 29). They were developed to simplify cell culture-based IAV vaccine production (28, 29). The siat7e cells are anchorage independent and were tested for their potentially more favorable compatibility to drop encapsulation (4, 7). IAV production was compared between the different cell lines to determine the kinetics of viral infection in a bulk infection. The cells were infected at an MOI of 0.1, and viral RNA was measured from supernatant at 0, 6, 12, 24, 30, 36, and 48 hpi using RT-qPCR, in which the 0 h measurement represents the inoculating dose (Fig. 2A to C). A549 cells demonstrated the most robust viral RNA production during

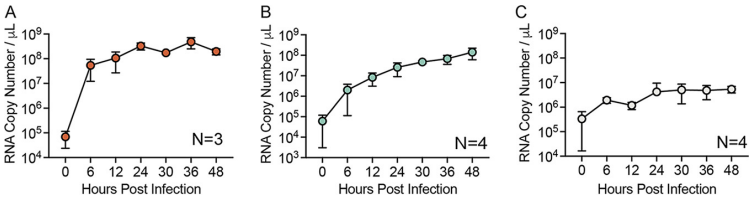


FIG 2 IAV infection in different cell lines. IAV genomes, as measured by RT-qPCR over 48 hpi from supernatant of infected A549 (A), MDCK (B), and siat7e (C) cells. All data represent the mean ± SD of a minimum of three independent replicates.

infection, with 9.5×10^4 genome copies/μL at 0 hpi, which increased to 3.9×10^8 genome copies/μL at 24 hpi, a 1,000-fold increase. Between 24 and 48 hpi, the amount of RNA detected fluctuated slightly but remained between 1.6 and 6.1×10^8 genome copies/μL. MDCK cells demonstrated a slower increase in viral production, with 1.1×10^5 genome copies/μL at 0 hpi, increasing to 3.5×10^7 genome copies/μL at 24 hpi, a 100-fold increase, before reaching 1.9×10^8 genome copies/μL at 48 hpi. IAV production in siat7e cells was limited with no exponential increase over the course of 48 hpi, with 6.0×10^5 genome copies/μL at 0 hpi and 5.4×10^6 genome copies/μL at 48 hpi. For A/California/07/2009, peak hemagglutinin (HA) titers in siat7e cells were previously observed at 24 hpi, although 50% tissue culture infective dose ($TCID_{50}$) measurements showed similar kinetics to our observed viral RNA measurements, with limited increases in titers from 0 to 120 hpi, suggesting that replication and production of infectious virus in the siat7e cells is limited (29). In addition, viral growth, as measured by $TCID_{50}$, of A/Brisbane/59/2007 IVR-148 (H1N1) and A/Uruguay/716/2007 X-175C (H3N2) in the siat7e cells was also not observed until 96 and 84 h, respectively, suggesting that there is variability in viral replication and kinetics in the siat7e cells (11). Overall, the average genomes/μL for A549, MDCK, and siat7e cells from 6 to 48 hpi, the time frame that represents active IAV production, was 2.6×10^8, 6.4×10^7, and 2.9×10^6 genome copies/μL, respectively, demonstrating cell type-dependent differences in IAV production.

Drop stability and cell viability during incubation. To perform long-term infection studies of viral infection in drops, drops must remain stable against coalescence, and cells must remain viable to propagate virus. Therefore, drop radii and cell viability over 24 h were first quantified before performing viral infection experiments. The A549, MDCK, and siat7e cells were encapsulated in 100-μm drops and incubated for 24 h. Drop radii (R) were measured following encapsulation at 0 h and after incubation of all three cell lines at 24 h (Fig. 3A). For A549 cells, the average R at 0 h was 45.1 μm with a 95% confidence interval (CI) of 35.9 to 54.3 μm, which decreased by 0.9 μm to 44.2 μm, with a 95% CI of 0.9 to 0.7 μm after incubation for 24 h. For MDCK cells, the average R at 0 h was 44.6 μm with a 95% CI of 35.3 to 53.9 μm, which increased by 1.3 μm to 45.9 μm with a 95% CI of 1.2 to 1.5 μm after incubation for 24 h. For the siat7e cells, the average R at 0 h was 44.8 μm with a 95% CI of 35.5 to 54.2 μm, which increased by 0.7 μm to 45.5 μm with a 95% CI of 0.6 to 0.9 μm after incubation for 24 h. A linear mixed-effects model was used to determine if there was a significant change in R as a function of the cell type or time point postencapsulation. The change in R from 0 to 24 h following encapsulation of either A549 or MDCK cells was considered significantly different ($P \leq 0.001$ for both), while the change in drop R from encapsulated siat7e cells was not significant ($P = 0.323$). However, since the changes in R for both cell lines are less than 3% of the original drop sizes and correspondingly less than the camera resolution of 1.6 μm/pixel, these data indicate that cell incubation does not have a meaningful impact on changes in drop stability as measured by R. Representative images of encapsulated A549 cells at 0 h (Fig. 3B) and 24 h (Fig. 3C) demonstrate this drop stability.

To assess cell viability, drops containing cells were broken after 24 h of incubation, and the aqueous layer which contained the cells was collected and analyzed with a colorimetric cell viability stain and by FACS analysis. Following staining and counts of the cells using a hemacytometer, each of the three cell lines showed cell viability over 75%. Average

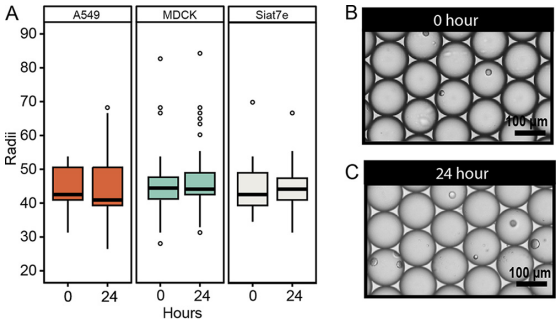


FIG 3 Drop stability in 100-μm microfluidic drops. (A) Drop radii (R) of encapsulated A549 (n = 4,067), MDCK (n = 4,235), and siat7e (n = 3,858) cells at 0 and 24 h. Representative images of A549 cells encapsulated in drops at 0 h (B) and 24 h (C). Scale bars, 100 μm.

live cell viability percentages were 76.2 $\pm$ 7.6% for A549 cells, 89.3 $\pm$ 4.6% for MDCK cells, and 90.6 $\pm$ 3.5% for siat7e cells. FACS analysis was used to quantify cell viability of a larger population of cells incubated in drops for 24 h (Fig. 4B to D). Cells were stained with propidium iodide to identify apoptotic cells within the population. A minimum of 10,000 cells were analyzed. The percentage of viable A549 cells following drop incubation was 83.1%. For MDCK cells, the percentage of viable cells was 90.7%, while viability of the siat7e cells was 94.5%. For all three cell lines, the percentage of viable cells determined by FACS was in the margin of error for the values determined by trypan blue staining.

The high cell viability observed in the siat7e cells was expected due to the cells being nonadherent with fewer physiological changes when encapsulated within drops. While the viability of the A549 cells was less than that observed in the MDCK and siat7e cells in drops, there was not a significant difference. We hypothesize that the differences observed in viability between the adherent A549 and MDCK cells were due to inherent differences between the cells themselves. MDCK cells are more prone to overgrowing and detaching from tissue culture flasks. They do not display strong contact inhibition and may be slightly better suited for drop encapsulation. In comparison, A549 cells demonstrate strong contact inhibition and rarely detach from tissue culture flasks. These inherent differences may impact cell viability following encapsulation into microfluidic drops.

Cell loading in microfluidic drops. For single-cell studies, each drop should be loaded with one cell. The number of cells loaded into drops can be estimated using a Poisson distribution (37) as follows:

$$p(k, \lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$$

where p is the probability of finding k cells per drop, k is the number of cells per drop, and λ is the mean number of cells in the volume of each drop. The λ is calculated by multiplying the starting cell density by the volume of a drop (22). Based upon Poisson statistics, loading cells at a density of 2 $\times$ 10^6 cells/mL into 100-μm drops should result in a λ of 1.0, with 37% of drops containing no cells, 37% of drops containing one cell, 18% of drops containing two cells, 6% of drops containing three cells, and 1.9% of drops containing four or more cells. To check if the three cell lines followed the predicted Poisson distribution, cells were loaded at a density of 2 $\times$ 10^6 cells/mL, and cell counts were made using still images of drops loaded onto a hemocytometer (Fig. 5A and B). Our calculated distribution did not follow expected Poisson statistics. The observed cell loading distributions for all three cell lines were much lower than expected. The percentages of drops that contained either 0, 1, 2, 3, or 4+ cells per drop were determined to understand how well the cell loading followed the predicted Poisson distribution. A549 cells had 10.8 $\pm$ 1.8% of total drops containing one cell from the images acquired on the hemocytometer, compared to the predicted 37% (Fig. 5C). MDCK cells had 8.0 $\pm$ 4.0% of total drops containing one cell (Fig. 5D), and siat7e cells had 5.7 $\pm$ 1.9% of

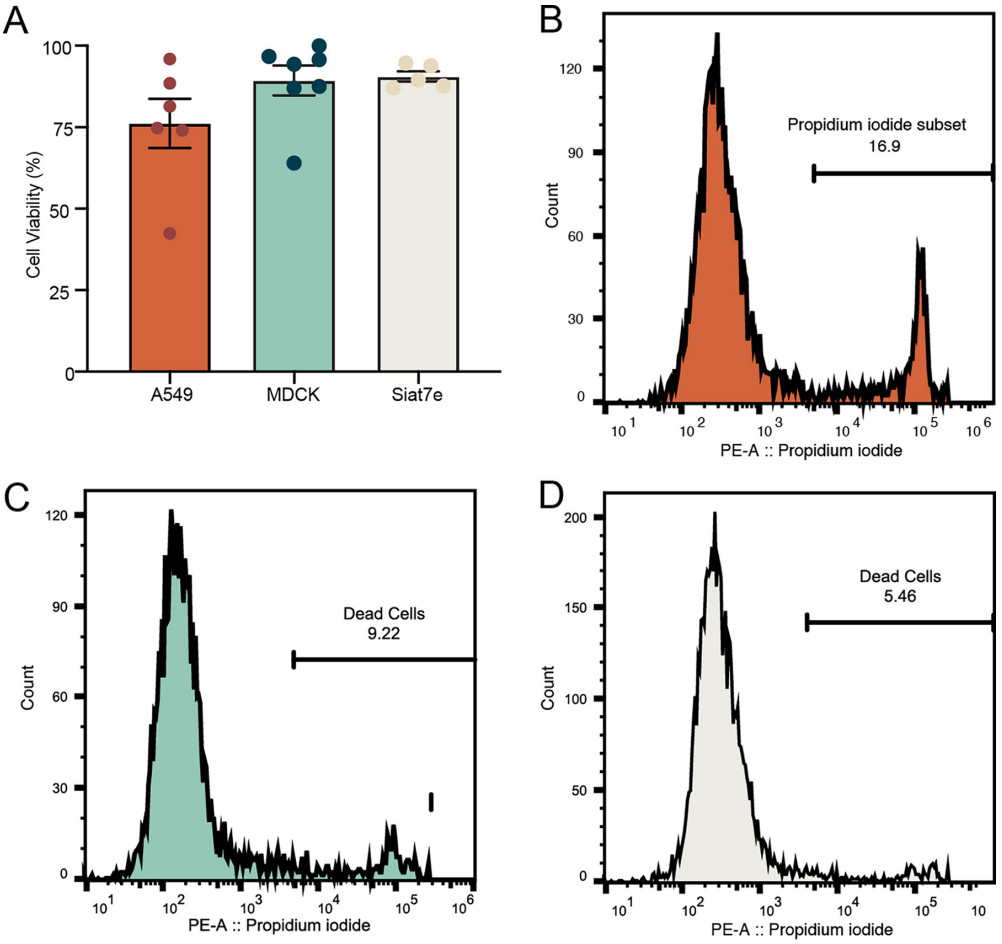


FIG 4 Cell viability in 100-μm microfluidic drops. (A) A549, MDCK, and siat7e cell viability at 24 h postencapsulation as determined by trypan blue staining, represented by the mean ± standard error of the mean (SEM). Points represent individual experiments. (B to D) FACS analysis of cell viability for A549 (B), MDCK (C), and siat7e (D) cells. The line on the histogram indicates the gating strategy to identify apoptotic cells.

total drops containing one cell on the hemocytometer (Fig. 5E). The percentage of loaded drops that contained a single cell only was also calculated. For A549 and MDCK cells, 68% and 73% of loaded drops contained a single cell, respectively (Table 1). For siat7e cells, only 38% of loaded drops contained a single cell (Table 1). With lower than expected cell loading data, we recalculated λ based on the experimentally determined loading distribution to accurately calculate the concentration of cells loaded into drops, not just added to the syringe for encapsulation. Based upon the experimental loading distribution, λ for each cell line was calculated using a Poisson mixed-effects model. The λ values correspond to starting cell concentrations that were lower than the input concentration of 2×10^6 cells/mL. The A549 cells had a λ of 0.282, which corresponds to an estimated cell concentration of 2.80×10^5 cells/mL. The MDCK cells had a λ of 0.164, which corresponds to an estimated cell concentration of 3.19×10^5 cells/mL. The siat7e cells had a λ of 0.382 and an estimated cell concentration of 3.85×10^5 cells/mL. We hypothesize that the discrepancy between the starting cell concentration and observed loading distribution was most likely due to cell settling in the syringe during loading and cell adherence to the filter upstream of the flow-focusing junction. The theoretical Poisson distributions that correspond to the calculated λ values for each cell line were determined (Fig. 5C to E). A chi-square analysis was used to evaluate how the actual

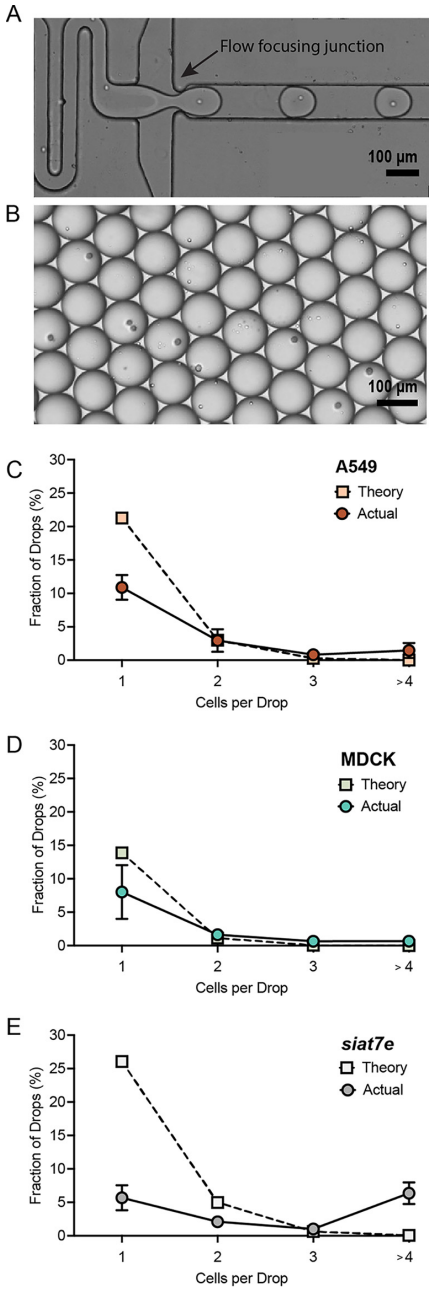


FIG 5 Cell encapsulation in microfluidic drops. (A) Still image of high-speed camera footage during cell encapsulation. The image has been modified to cover lithography marks outside the channels for aesthetics. (B) Representative image of drops encapsulated onto a hemocytometer. (C) A549 cell loading as measured with the hemocytometer (red circle, n = 600) compared to the theoretical Poisson distribution (peach square). (D) MDCK cell loading as measured with the hemocytometer (aqua circle, n = 601) compared to the theoretical Poisson distribution (green square). (E) Siat7e cell loading as measured with the hemocytometer (light gray circle, n = 598) and compared to the theoretical Poisson distribution (white square). Data points are represented by the mean ± SD. All scale bars are 100 μm.

TABLE 1 Percentage of cells/drop in loaded drops containing cells

Cell type	% 1 cell/drop	% 2 cells/drop	% 3 cells/drop	% 4+ cells/drop
A549	67.6	18.2	5.1	9.1
MDCK	72.9	14.9	6.0	6.2
siat7e	36.5	15.5	6.2	41.8

distribution of cells in drops compared to the theoretical Poisson distributions. A higher *P* value indicates a lack of evidence that the measured distribution does not follow a Poisson distribution. The A549 cells had the highest *P* value of 0.70, while the MDCK cells had a *P* value of 0.48. The drops containing the siat7e cells had the lowest *P* value of 0.15, suggesting that the distribution of siat7e cells in drops was the least representative of a Poisson distribution out of the three cell lines. Overall, the siat7e cells demonstrated the greatest variability and unpredictability in cell loading. We hypothesize that this is due to the cells adhering together, which is also visible during normal cell growth in a shaker flask in which large clumps of cells arise during growth. Due to poor cell encapsulation and low IAV replication during infection, siat7e cells were excluded from further analysis.

IAV propagation in microfluidic drops. Following analysis of cell encapsulation and viability of A549 and MDCK cells, IAV propagation was compared between A549 and MDCK cells in bulk and in drops. A549 and MDCK cells were infected at an MOI of 0.1 with the A/California/07/2009 (H1N1) IAV strain. To quantify the increase in virus genomes and infectious progeny for both bulk and drop infections at 0 and 24 hpi, viral RNA was measured using RT-qPCR, and infectious virus output was measured using plaque assay. Measurements were taken at 0 and 24 hpi to limit detection of secondary spread within bulk cultures, as a single round of replication occurs between 16 and 24 hpi (38). For A549 cells, the number of genome copies/μL in drops, as measured by M gene RNA copy number using RT-qPCR, was 2.3×10^5 genome copies/μL at 0 h and increased to 2.3×10^7 genome copies/μL at 24 h (Fig. 6A). Bulk infections of A549 cells had a comparable increase, with 1.8×10^6 genome copies at 0 h and 5.6×10^8 genome copies/μL at 24 h (Fig. 6A). The log RNA concentrations in drops compared to bulk at 0 h and 24 h were significantly different ($P = 0.0006$ and $P = 4.3E-09$, respectively), which we hypothesize is due to a lower number of total cells loaded into drops during encapsulation as described in "Cell loading in microfluidic drops" than cells seeded on tissue culture plates. While this was not seen in the analysis of infectious virus (PFU/mL) for A549 cells ($P = 0.31$ for 0 h and $P = 0.71$ for 24 h) (Fig. 6B), it was observed for both RNA ($P = 7.2E-05$ for 0 h and $P = 0.0002$ for 24 h) and PFU/mL for MDCK cells (Fig. 6C and D). However, the log difference of viral genomes produced by cells from 0 to 24 h in drops compared to cells in bulk was not significant ($P = 0.057$) and suggests that overall virus production was not impacted by cell encapsulation. Recovery of infectious virus from A549 cells over the same 24-h incubation period was not significantly different ($P = 0.84$) between IAV infection in drops and bulk culture with 1.1×10^5 PFU/mL and 3.6×10^5 PFU/mL recovered at 24 hpi, respectively (Fig. 6B).

For MDCK cells, the number of genome copies in drops was 8.9×10^3 genome copies/μL at 0 h and increased to 2.3×10^7 genome copies/μL at 24 h, whereas bulk infections of MDCK cells increased from 1.2×10^5 genome copies/μL at 0 h and 3.4×10^8 genome copies/μL at 24 h (Fig. 6C). Similar to A549 cells, the overall log difference of viral genomes produced by cells from 0 to 24 h was not significantly different ($P = 0.6$) between drop and bulk infection. Recovery of infectious virus from MDCK cells over 24 h in drops was 2.6×10^3 PFU/mL, while bulk infections produced 1.0×10^6 PFU/mL (Fig. 6D). In contrast to A549 cells, the log PFU/mL concentrations measured in MDCK cells in drops compared to bulk at 0 h and 24 h were significantly different ($P = 0.0004$ and $P = 0.003$, respectively); however, this did not translate into a significant difference in the change in log PFU/mL concentration from 0 to 24 h between drop and bulk infections. The PFU-to-genome ratio for MDCK cells was 1 PFU per 5.81×10^3 genomes in drops and 1 PFU per 3.24×10^2 genomes in bulk. In comparison, for A549 cells, the PFU-to-genome ratio results in 1 PFU per 2.97×10^2 genomes in drops and 1 PFU per 1.63×10^3 genomes in bulk.

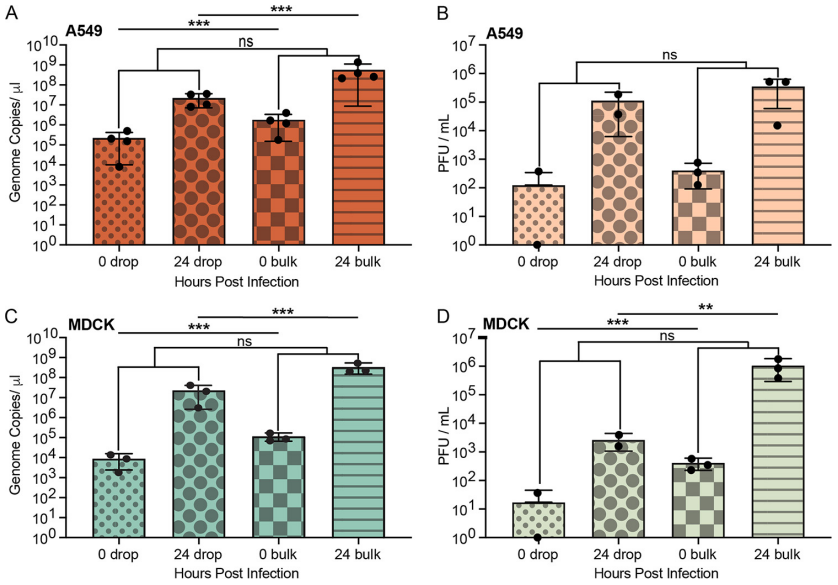


FIG 6 Comparison of drop and bulk IAV infections in A549 and MDCK cells. (A) RNA copy number of A/California/07/2009 (H1N1) IAV in A549 cells infected at an MOI of 0.1 at 0 hpi in drops (red small dots), 24 hpi in drops (red large dots), 0 hpi in bulk (red squares), and 24 hpi in bulk (red stripe). (B) Infectious virions (PFU/mL) released from A549 cells at 0 hpi in drops (peach small dots), 24 hpi in drops (peach large dots), 0 hpi in bulk (peach squares), and 24 hpi in bulk (peach stripe). (C) RNA copy number of A/California/07/2009 (H1N1) IAV in MDCK cells infected at an MOI of 0.1 at 0 hpi in drops (aqua small dots), 24 hpi in drops (aqua large dots), 0 hpi in bulk (aqua squares), and 24 hpi in bulk (aqua stripe). (D) Infectious virions (PFU/mL) released from MDCK cells 0 hpi in drops (green small dots), 24 hpi in drops (green large dots), 0 hpi in bulk (green squares), and 24 hpi in bulk (green stripe). All data represented as mean ± the SD with a minimum of three independent experiments. Significance is based upon a two-tailed *t* test (**, $P < 0.01$; ***, $P < 0.001$; ns, not significant).

DISCUSSION

The drop-based microfluidic methods developed in this work will enable future studies of single-cell IAV infections using multiple cell lines. The methods expand the application of drop-based microfluidics from the nonenveloped, positive-sense MNV-1 (4, 7, 8) to include the enveloped and negative-sense IAV, thus broadening the scope of single-cell virology assays. Additionally, this work can serve as a blueprint for testing, comparing, and implementing drop-based microfluidics methods in future single-cell studies. With the recent interest and development of single-cell omic technologies for studying viral infections (13–15, 20, 30–36), we expect this work to further expand capabilities and increase applications of single-cell technologies in virology.

To successfully perform in drop infections, standard cell lines used for propagating viruses need to remain viable, and this depends upon both drop size and cell type. Previous studies exploring viability of cells in drops found that smaller drop sizes resulted in low levels of cell viability, even after a couple of hours, which is most likely due to a lack of enough nutrients or buildup of waste products within the drops (39). Here, we demonstrate that 100-μm-diameter drops yielded high cell viability and drop stability following 24 h of incubation. Cell viability of A549, MDCK, and siat7e cells in 100-μm drops remained high (>75%), and the average drop radii changed by less than 3%. High cell viability was expected from the siat7e suspension cell line, which was found to have a mean of 91%. Though more variable between experimental replicates compared to the siat7e cells, the adherent A549 and MDCK cell lines still demonstrated high cell viability (Fig. 3D).

To analyze the distribution of cells in drops, we used a hemocytometer to determine cell loading distributions. A concentration of 2×10^6 cells/mL was calculated to obtain a maximum number of 100-μm diameter drops that contain a single cell based upon Poisson loading. However, throughout our experiments, cells settled in the

syringes and attached to the aqueous-stream filter on the microfluidic drop maker, lowering the number of cells being encapsulated. Therefore, the average number of cells per drop volume, as described by λ in our experiments, corresponded to a lower cell concentration on the order of 10^5 cells/mL. One option to reduce cell settling is the use of density-matched medium cells during encapsulation with a biologically inert polymer, such as OptiPrep (40). This would decrease the settling velocity of a cell as estimated using Stokes' law (40). However, cell media components would have to be further optimized, as increases in OptiPrep concentration also decrease cell viability (40). Alternatively, a larger vessel in which cells are constantly stirred and then pneumatically transferred into the microfluidic device would also decrease cell settling. Finally, cell adherence to the upstream filter was caused by clumps of cells, likely free DNA and cell debris released during cell lysis. A reduction in shear stress that the cells experience by using gentle pipetting or syringe loading would aid in reducing cell debris. Cell adherence to device structures can be reduced by eliminating the filter structures or increasing the gaps between the pillars within the microfluidic device.

While cell settling impacted the cell loading concentration, the majority of our loaded drops still contained a single cell. Surprisingly, the suspension siat7e cell line had the largest variation from the expected Poisson distribution, with more drops having four or more cells/drop than was observed for A549 or MDCK cells. We hypothesize that this large variation was due to the siat7e cells adhering to each other when cultured in a spinner flask and during infection prior to cell loading. To improve cell loading, a higher concentration of cells can be used, although previous analysis of 2×10^6, 4×10^6, and 8×10^6 cells/mL in 50-μm drops had similar loading distributions (4). Thus, we conclude that cell loading depends on the cell type, with adherent A549 and MDCK cells resulting in a majority of single cells loaded into drops compared to the suspension siat7e cells.

Following our success in encapsulation and culturing multiple cell lines within microfluidic drops, we investigated whether infected cells produce virus following encapsulation compared to bulk studies. Due to the highly variable cell loading and low virus production observed in the siat7e cells, only the A549 and MDCK cells were utilized for drop infection experiments. Measurements of viral RNA via RT-qPCR and infectious progeny via plaque assay were made at 0 and 24 h from supernatant collection from broken drops and compared to standard bulk infections on well plates. The amount of virus produced from A549 and MDCK cells between 0 and 24 h in bulk was not significantly different than the amount of virus produced from cells in drops. This demonstrates that microfluidic drops can be used to propagate IAV from individual cells and that standard adherent cell lines can be used without the need to adapt adherent cells to suspension cultures. To further evaluate single-cell kinetics of infection, multiple time points could be analyzed from infected cells in drops. One proposed method for performing this analysis is the use of fluorescent protein (FP) expression from specific viral or cellular gene targets. This would allow for continuous observations of cells in drops over time while quantifying changes in FP expression that correlate with virus replication or transcription. Time-resolved control using drops could expand the capabilities of previous work in which fluorescent proteins expressed by herpes simplex virus 1 were used to isolate infected cells at 5 hpi for downstream single-cell RNA sequencing (scRNA-seq) (14). This work lays the foundation to expand methods for investigating the heterogeneity within population dynamics at the single-cell level.

One concern regarding incubating viruses in drops is the transfer of viruses between drops and the potential loss of infectious virus upon drop breaking. The surfactant that is used to stabilize the drops is a mixture of di-block and tri-block polymers with fluorinated tails that extend into the oil phase and hydrophilic blocks that extend into the aqueous phase that are PEGylated (26). Polyethylene glycol (PEG) prevents most biological materials from interacting with the drop interface and diffusing between neighboring drops. Small molecules (water, amino acids, ions, and some antibiotics) have been known to diffuse between neighboring drops (41–43); however, under our drop conditions, we empirically observe that viruses do not cross the drop interface.

Conclusions. Analysis of individually infected cells has revealed the complexity and heterogeneity of viral infections (9–11, 20). The implementation of current omics technologies in these systems has demonstrated that viral infections are as heterogeneous and complex at the single-cell level as they are at the systems level (23–25). The application of new technologies for the study of virology can elucidate how individual host cells influence viral evolution and respond to infection through the innate and adaptive immune systems. Considering the current COVID-19 pandemic and the emergence of multiple variants within the viral population, it is critical to expand our ability to study of how these variants arise, how they operate under selective pressures, and how they impact transmission and virulence (44). Single-cell studies can be limited by low-throughput analysis, reliance on expensive cell-sorting machinery, and genetic manipulation of the virus itself. The use of drop-based microfluidics can expand single-cell studies, offering the ability to perform high-throughput analysis of individually infected cells (5, 21, 31). Advantageously, these applications can be adapted to many laboratories with minimal investment. The ability to expand single-cell analysis using drop-based microfluidics has the potential to revolutionize the field of virology.

MATERIALS AND METHODS

Cells and viruses. Human alveolar epithelial cells (A549), Madin-Darby canine kidney cells (MDCK), and MDCK plus human siat7e gene (siat7e) cells were obtained from ATCC (CCL-185, CCL-34) and Joseph Shiloach (NIH), respectively (28, 29). A549 cells were propagated in Hams F-12 (Corning) medium supplemented with 10% fetal bovine serum (HyClone) and 1× penicillin-streptomycin (Fisher Scientific). MDCK and siat7e cells were propagated in Dulbecco's modified Eagle medium (DMEM; Corning) supplemented with 10% fetal bovine serum (HyClone) and 1× penicillin-streptomycin. The IAV virus strain A/California/07/2009 (H1N1) was obtained from Christopher Brooke (University of Illinois Urbana-Champaign [UIUC]). Stocks of A/California/07/2009 were propagated, and we determined their titer on MDCK cells.

Microfluidic device fabrication. Microfluidic devices for drop making were fabricated by patterning SU-8 photoresist (Microchem; catalog no. SU-8 50) on silicon wafers (University Wafer; ID no. 447; test grade) to create 100-μm-tall and 100-μm-wide channels. Polydimethylsiloxane (PDMS) (Sylgard 184) was poured onto the wafers at a 10:1 mass ratio of polymer to cross-linking agent according to standard photolithography techniques (45). Air was purged from the uncured PDMS by placing the filled mold in a vacuum chamber for at least 1 h. The PDMS was cured in an oven at 55°C for 24 h, and ports were punched into the PDMS slab with a 0.75-mm-inside-diameter biopsy specimen punch (EMS Rapid-Core; Electron Microscopy Sciences). The PDMS slab was bonded to a 2-in. by 3-in. glass slide (VWR; microslides, catalog no. 48382-179) after plasma treatment (Harrick Plasma; catalog no. PDC-001) for 60 s at high power (30 W) and 700 mtorr oxygen pressure. The drop-making devices were made hydrophobic by flowing Aquapel (Pittsburgh Glass Works) through the channels, followed by blowing the channels with air filtered through a GVS Abluo 25-mm 0.2-μm filter (Fisher Scientific) before baking the devices in an oven at 55°C for 1 h.

Cell encapsulation in drops. Cells were seeded onto T25 flasks or 6-well plates at a density of 1×10^5 cells/cm^2 and incubated overnight. Cells were prepared for encapsulation by washing 2× with phosphate-buffered saline (PBS) followed by trypsinization to remove cells from the surface of the flasks or plates. The appropriate culture media containing 10% fetal bovine serum (FBS) were added to the cells to neutralize trypsinization, and cells were collected and centrifuged at 500 × *g* for 5 min. The cell pellet was gently washed with 3 to 5 mL of PBS and centrifuged a second time at 500 × *g* for 5 min. The cell pellets were gently resuspended in FBS-free media to reach 2×10^6 cells/mL in their designated media. We placed 10 μL of cells on a hemocytometer to visualize and confirm cell concentration and verify the presence of mostly single cells. Cells were loaded into a 3-mL syringe (BD) for injection onto a flow-focusing microfluidic device for encapsulation into 100-μm drops (~523 pL) (46). Drops were stabilized in a fluorinated HFE7500 oil (3 M) continuous phase with a 1.5% (wt/wt) solution of PEG-PFPE$_2$-based surfactant (RAN Biotechnologies; 008-FluoroSurfactant). The fluids were transferred into the microfluidic device with syringe pumps (New Era NE-1000) controlled by a custom LabVIEW (2015) program at flow rates of 1,000 μL/h for the aqueous phase and 3,000 μL/h for the oil phase. A 1,500-μL total volume consisting of 375 μL of drops containing cells and 1,125 μL of oil was collected in a 2-mL microcentrifuge tube (Eppendorf). Following encapsulation, drops containing cells were incubated in an open microcentrifuge tube covered with parafilm at 37°C with 95% relative humidity and 5% CO_2. Cells were released from drops by removing most of the oil phase and then adding 1 mL of a 20% wt/wt 1H,1H,2H,2H-perfluoro-1-octanol (PFO) in HFE7500 oil to break the emulsion, followed by vortexing. Centrifugation at 500 × *g* for 5 min was used to separate the aqueous cell suspension from the oil phase to isolate cells for further analysis.

Drop measurements. Drop radii were measured after imaging a drop monolayer on a hemocytometer with a charge-coupled-device (CCD) camera (Flir Grasshopper3) on an inverted microscope (Nikon Ti-U). A monolayer of drops with minimal packing density is viewed using a hemocytometer so that the drops remained spherical for image analysis. A custom MATLAB script was used to determine the drop radii, *R*, at 0 h and 24 h after the cells were incubated. Six images of drops, with a minimum of 300 drops per image per time point, were analyzed. A data set containing *R* values at 0 and 24 h post encapsulation, the encapsulated cell line, and experimental dates was created to analyze changes in *R*. A linear mixed-effects model, with the experimental date as a random factor, was used to analyze the contributions of cell type and time point to the data set.

Cell loading and viability. The distribution of cells per drop was determined using still images of a drop monolayer on a hemocytometer as described in "Drop measurements." These distributions were fit to a

Poisson mixed-effects model with cell lines as the fixed effect and date as the random effect (47, 48). We used a Pearson chi-square goodness-of-fit test to determine whether the observed data did not fit the theoretical distribution. Following encapsulation, drops were incubated at 37°C with 95% relative humidity and 5% CO_2. Cells were released from drops as previously described. The cells in the aqueous phase were collected and pelleted at 1,500 $\times$ *g* for 5 min. The cell pellet was resuspended in 200 μL PBS. Cells were diluted 1:5 in PBS with 10% trypan blue stain for a final volume of 50 μL to determine cell viability. An unpaired *t* test was used to determine if cell viability was significantly different between the cell lines. Flow cytometry (FACS) analysis was performed on cells isolated as described above and resuspended in PBS with 2% FBS and 5 μg/mL propidium iodide (Sigma-Aldrich; catalog no. 537060) to stain apoptotic cells. Samples were analyzed with a BD LSRFortessa flow cytometer (BD Biosciences) and FlowJo software (Tree Star).

IAV infection of adherent cells (A549 and MDCK). Cells were seeded onto a 6-well plate at a concentration of 1 $\times$ 10^6 cells/well and infected with IAV H1N1 at an MOI of 0.1 based upon plaque assay titering in infection media. The infection media consisted of Hams F-12 or DMEM supplemented with 1 mM HEPES (HyClone), 1$\times$ penicillin-streptomycin, 0.1% bovine serum albumin (BSA) (MP Biomedical), and 1 μg/mL of tosylsulfonyl phenylalanyl chloromethyl ketone (TPCK)-trypsin (Worthington Biomedical) (49). Briefly, cells were washed with 1$\times$ PBS (Corning) and incubated with 200 μL of virus inoculum for 1 h. The inoculum was removed and replaced with 1.5 mL fresh infection medium for another 1 h of incubation. Infection medium was removed, cells were washed with PBS, and the infected cells were detached from the plate by trypsinization. Infected cells were processed as described in "Cell encapsulation in drops" and resuspended in encapsulation media containing either Hams F-12 or DMEM and 1 mM HEPES, 1$\times$ penicillin-streptomycin, and 0.1% BSA. The infected cell suspension was split into two populations, with one population replated as a bulk control and the other encapsulated in 100-μm drops. Bulk and drop infections were incubated at 37°C with 95% relative humidity and 5% CO_2 and frozen at −20°C at 0 and 24 h postinfection (hpi).

IAV infection of suspension siat7e cells. To infect the siat7e suspension cells at an MOI of 0.1 based upon plaque assay titering, 1 $\times$ 10^7 cells were pelleted at 500 $\times$ *g* for 5 min in a 50-mL conical centrifuge tube and then resuspended in 200 μL of virus inoculum in DMEM supplemented with 1 mM HEPES (HyClone), 1$\times$ penicillin-streptomycin, 0.1% BSA, and 1 μg/mL of TPCK-trypsin (49). The 50-mL tube containing the cells and virus inoculum was shaken at 90 rpm for 1 h. The cells were pelleted at 500 $\times$ *g* for 5 min, and the inoculum was removed and replaced with fresh DMEM infection media (10 mL) for bulk infections or DMEM encapsulation media as described in "IAV infection of adherent cells (A549 and MDCK)." Infected cells were either encapsulated in 100-μm drops or replated as a bulk control. For bulk infections, a 1-mL volume of cells was added to each well of a 6-well plate (2 total) and placed on the shaker. Encapsulated cells were processed as described in "Cell encapsulation in drops." Bulk and drop infections were incubated at 37°C with 95% relative humidity and 5% CO_2 and frozen at −20°C at 0 and 24 hpi.

M gene abundance by RT-qPCR. RNA was extracted from cell suspensions using the Qiagen QIAamp viral RNA minikit. RNA copy number of the IAV matrix protein gene (M gene) was determined using a TaqMan RT-qPCR assay (5). Amplification primer sequence was originally described by Shu et al. (50) as follows: M gene forward, 5′-GACCRATCCTGTCACCTCTGAC-3′, and M gene reverse, 5′-AGGGCATTCTGGACAAATCGTCTA-3′. The sequence of the M gene TaqMan probe was 5′-/FAM/TGCAGTCCTCGCTCACTGGGCACG/BHQ1/-3′. Working stocks of the primers and probe (Eurofins Operon) were prepared at 25 μM and 10 μM, respectively, for use in the RT-qPCR. Samples were amplified using a SuperScript III Platinum one-step RT-qPCR kit (Invitrogen; catalog no. 11732-020) with a final reaction volume of 25 μL. Each reaction mix contained 400 nM M gene forward and reverse primers, 200 nM M gene TaqMan probe, 0.05 μM ROX reference dye, 5 U/μL Superase RNase inhibitor (Invitrogen; catalog no. AM2694), and 2.5 μL of RNA template. Thermocycling was performed in an RT-qPCR machine (QuantStudio 3; Applied Biosystems) with the following cycling conditions: 1 cycle for 30 min at 60°C, 1 cycle for 2 min at 95°C, and 40 cycles between 15 s at 95°C and 1 min at 60°C.

Plaque assay. Postinfection IAV titers were determined by plaque assay on MDCK cells seeded in 6-well plates at 1 $\times$ 10^6 cells per well. For drop infections, the viral supernatants were sampled from broken drops at 0 and 24 hpi. For bulk infections, viral supernatants were sampled from the tissue culture plate wells at 0 and 24 hpi. The supernatants were serially diluted in DMEM media with 1 mM HEPES (HyClone), 1$\times$ penicillin-streptomycin, 0.1% BSA, and 2 μg/mL of TPCK-trypsin. Overlay medium consisting of 2$\times$ MEM with 2$\times$ penicillin-streptomycin, 1 mM HEPES, 2 μg/mL of TPCK, and 3% carboxymethyl cellulose (MP Biomedical) with 0.2 mg/mL DEAE-dextran (MP Biomedical) was added, and well plates were incubated at 37°C for 4 days. Plaques were fixed with 10% buffered formalin (Fisher), washed with deionized (DI) water, and stained with 0.5% crystal violet (Thermo Fisher Scientific) for visualization.

ACKNOWLEDGMENTS

This work was supported by the Defense Advanced Research Projects Agency (DARPA) grant W911NF-17-2-0034, National Institute of Allergy and Infectious Diseases (NIAID) of the National Institutes of Health (NIH) under award number 1R56AI156137, and National Science Foundation (NSF) CAREER grant 1753352.

We thank Chris Brooke at the University of Illinois at Urbana-Champaign for providing A/California/07/2009 (H1N1) and human alveolar epithelial A549 cells, Joseph Shiloach at the NIH for providing siat7e cells, Agnieszka Rynda-Apple at Montana State University for providing MDCK cells, and Albert Parker at Montana State University for assistance with data analysis and statistics.

REFERENCES

1. Dolan PT, Whitfield ZJ, Andino R. 2018. Mapping the evolutionary potential of RNA viruses. Cell Host Microbe 23:435–446. https://doi.org/10.1016/j.chom.2018.03.012.
2. Petrova VN, Russell CA. 2018. The evolution of seasonal influenza viruses. Nat Rev Microbiol 16:47–60. https://doi.org/10.1038/nrmicro.2017.118.
3. Brooke CB. 2017. Population diversity and collective interactions during influenza virus infection. J Virol 91:e01164-17. https://doi.org/10.1128/JVI.01164-17.
4. Fischer AE, Wu SK, Proescher JBG, Rotem A, Chang CB, Zhang H, Tao Y, Mehoke TS, Thielen PM, Kolawole AO, Smith TJ, Wobus CE, Weitz DA, Lin JS, Feldman AB, Wolfe JT. 2015. A high-throughput drop microfluidic system for virus culture and analysis. J Virol Methods 213:111–117. https://doi.org/10.1016/j.jviromet.2014.12.003.
5. Loveday EK, Zath GK, Bikos DA, Jay ZJ, Chang CB. 2021. Screening of additive formulations enables off-chip drop reverse transcription quantitative polymerase chain reaction of single influenza A virus genomes. Anal Chem 93:4365–4373. https://doi.org/10.1021/acs.analchem.0c03455.
6. Rotem A, Serohijos AWR, Chang CB, Wolfe JT, Fischer AE, Mehoke TS, Zhang H, Tao Y, Lloyd Ung W, Choi J-M, Rodrigues JV, Kolawole AO, Koehler SA, Wu S, Thielen PM, Cui N, Demirev PA, Giacobbi NS, Julian TR, Schwab K, Lin JS, Smith TJ, Pipas JM, Wobus CE, Feldman AB, Weitz DA, Shakhnovich EI. 2018. Evolution on the biophysical fitness landscape of an RNA virus. Mol Biol Evol 35:2390–2400. https://doi.org/10.1093/molbev/msy131.
7. Tao Y, Rotem A, Zhang H, Chang CB, Basu A, Kolawole AO, Koehler S, Ren Y, Lin JS, Pipas JM, Feldman AB, Wobus C, Weitz DA. 2015. Rapid, targeted and culture-free viral infectivity assay in drop-based microfluidics. Lab Chip 15:3934–3940. https://doi.org/10.1039/C5LC00556F.
8. Tao Y, Rotem A, Zhang H, Cockrell SK, Koehler SA, Chang CB, Ung LW, Cantalupo PG, Ren Y, Lin JS, Feldman AB, Wobus CE, Pipas JM, Weitz DA. 2015. Artifact-free quantification and sequencing of rare recombinant viruses by using drop-based microfluidics. Chembiochem 16:2167–2171. https://doi.org/10.1002/cbic.201500384.
9. Russell AB, Elshina E, Kowalsky JR, te Velthuis AJ, Bloom JD. 2019. Single-cell virus sequencing of influenza infections that trigger innate immunity. J Virol 93:e00500-19. https://doi.org/10.1128/JVI.00500-19.
10. Russell AB, Trapnell C, Bloom JD. 2018. Extreme heterogeneity of influenza virus infection in single cells. Elife 7:e32303. https://doi.org/10.7554/eLife.32303.
11. Sun J, Vera JC, Drnevich J, Lin YT, Ke R, Brooke CB. 2020. Single cell heterogeneity in influenza A virus gene expression shapes the innate antiviral response to infection. PLoS Pathog 16:e1008671. https://doi.org/10.1371/journal.ppat.1008671.
12. Heldt FS, Kupke SY, Dorl S, Reichl U, Frensing T. 2015. Single-cell analysis and stochastic modelling unveil large cell-to-cell variability in influenza A virus infection. Nat Commun 6:8938. https://doi.org/10.1038/ncomms9938.
13. Kupke SY, Ly L-H, Börno ST, Ruff A, Timmermann B, Vingron M, Haas S, Reichl U. 2020. Single-cell analysis uncovers a vast diversity in intracellular viral defective interfering RNA content affecting the large cell-to-cell heterogeneity in influenza A virus replication. Viruses 12:71. https://doi.org/10.3390/v12010071.
14. Drayman N, Patel P, Vistain L, Tay S. 2019. HSV-1 single-cell analysis reveals the activation of anti-viral and developmental programs in distinct sub-populations. Elife 8:e46339. https://doi.org/10.7554/eLife.46339.
15. Guo F, Li S, Caglar MU, Mao Z, Liu W, Woodman A, Arnold JJ, Wilke CO, Huang TJ, Cameron CE. 2017. Single-cell virology: on-chip investigation of viral infection dynamics. Cell Rep 21:1692–1704. https://doi.org/10.1016/j.celrep.2017.10.051.
16. Timm A, Yin J. 2012. Kinetics of virus production from single cells. Virology 424:11–17. https://doi.org/10.1016/j.virol.2011.12.005.
17. Zhu Y, Yongky A, Yin J. 2009. Growth of an RNA virus in single cells reveals a broad fitness distribution. Virology 385:39–46. https://doi.org/10.1016/j.virol.2008.10.031.
18. Schulte MB, Andino R. 2014. Single-cell analysis uncovers extensive biological noise in poliovirus replication. J Virol 88:6205–6212. https://doi.org/10.1128/JVI.03539-13.
19. Akpinar F, Timm A, Yin J. 2016. High-throughput single-cell kinetics of virus infections in the presence of defective interfering particles. J Virol 90:1599–1612. https://doi.org/10.1128/JVI.02190-15.
20. Combe M, Garijo R, Geller R, Cuevas JM, Sanjuán R. 2015. Single-cell analysis of RNA virus infection identifies multiple genetically diverse viral genomes within single infectious units. Cell Host Microbe 18:424–432. https://doi.org/10.1016/j.chom.2015.09.009.
21. Matuła K, Rivello F, Huck WTS. 2020. Single-cell analysis using droplet microfluidics. Adv Biosyst 4:e1900188. https://doi.org/10.1002/adbi.201900188.
22. Mazutis L, Gilbert J, Ung WL, Weitz DA, Griffiths AD, Heyman JA. 2013. Single-cell analysis and sorting using droplet-based microfluidics. Nat Protoc 8:870–891. https://doi.org/10.1038/nprot.2013.046.
23. Xu X, Wang J, Wu L, Guo J, Song Y, Tian T, Wang W, Zhu Z, Yang C. 2020. Microfluidic single-cell omics analysis. Small 16:e1903905. https://doi.org/10.1002/smll.201903905.
24. Prakadan SM, Shalek AK, Weitz DA. 2017. Scaling by shrinking: empowering single-cell 'omics' with microfluidic devices. Nat Rev Genet 18:345–361. https://doi.org/10.1038/nrg.2017.15.
25. Zilionis R, Nainys J, Veres A, Savova V, Zemmour D, Klein AM, Mazutis L. 2017. Single-cell barcoding and sequencing using droplet microfluidics. Nat Protoc 12:44–73. https://doi.org/10.1038/nprot.2016.154.
26. Holtze C, Rowat AC, Agresti JJ, Hutchison JB, Angile FE, Schmitz CH, Koster S, Duan H, Humphry KJ, Scanga RA, Johnson JS, Pisignano D, Weitz DA. 2008. Biocompatible surfactants for water-in-fluorocarbon emulsions. Lab Chip 8:1632–1639. https://doi.org/10.1039/b806706f.
27. Henderson KS. 2008. Murine norovirus, a recently discovered and highly prevalent viral agent of mice. Lab Anim (NY) 37:314–320. https://doi.org/10.1038/laban0708-314.
28. Chu C, Lugovtsev V, Golding H, Betenbaugh M, Shiloach J. 2009. Conversion of MDCK cell line to suspension culture by transfecting with human siat7e gene and its application for influenza virus production. Proc Natl Acad Sci U S A 106:14802–14807. https://doi.org/10.1073/pnas.0905912106.
29. Chu C, Lugovtsev V, Lewis A, Betenbaugh M, Shiloach J. 2010. Production and antigenic properties of influenza virus from suspension MDCK-siat7e cells in a bench-scale bioreactor. Vaccine 28:7193–7201. https://doi.org/10.1016/j.vaccine.2010.08.059.
30. Cristinelli S, Ciuffi A. 2018. The use of single-cell RNA-Seq to understand virus–host interactions. Curr Opin Virol 29:39–50. https://doi.org/10.1016/j.coviro.2018.03.001.
31. Gérard A, Woolfe A, Mottet G, Reichen M, Castrillon C, Menrath V, Ellouze S, Poitou A, Doineau R, Briseno-Roa L, Canales-Herrerias P, Mary P, Rose G, Ortega C, Delincé M, Essono S, Jia B, Iannascoli B, Richard-Le Goff O, Kumar R, Stewart SN, Pousse Y, Shen B, Grosselin K, Saudemont B, Sautel-Caillé A, Godina A, McNamara S, Eyer K, Millot GA, Baudry J, England P, Nizak C, Jensen A, Griffiths AD, Bruhns P, Brenan C. 2020. High-throughput single-cell activity-based screening and sequencing of antibodies using droplet microfluidics. Nat Biotechnol 38:715–721. https://doi.org/10.1038/s41587-020-0466-7.
32. Liao M, Liu Y, Yuan J, Wen Y, Xu G, Zhao J, Cheng L, Li J, Wang X, Wang F, Liu L, Amit I, Zhang S, Zhang Z. 2020. Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. Nat Med 26:842–844. https://doi.org/10.1038/s41591-020-0901-9.
33. Lin J, Jordi C, Son M, Van Phan H, Drayman N, Abasiyanik MF, Vistain L, Tu H-L, Tay S. 2019. Ultra-sensitive digital quantification of proteins and mRNA in single cells. Nat Commun 10:3544. https://doi.org/10.1038/s41467-019-11531-z.
34. Sen A, Rothenberg ME, Mukherjee G, Feng N, Kalisky T, Nair N, Johnstone IM, Clarke MF, Greenberg HB. 2012. Innate immune response to homologous rotavirus infection in the small intestinal villous epithelium at single-cell resolution. Proc Natl Acad Sci U S A 109:20667–20672. https://doi.org/10.1073/pnas.1212188109.
35. Timm AC, Warrick JW, Yin J. 2017. Quantitative profiling of innate immune activation by viral infection in single cells. Integr Biol (Camb) 9:782–791. https://doi.org/10.1039/c7ib00082k.
36. Zanini F, Pu S-Y, Bekerman E, Einav S, Quake SR. 2018. Single-cell transcriptional dynamics of flavivirus infection. Elife 7:e32942. https://doi.org/10.7554/eLife.32942.
37. Collins DJ, Neild A, deMello A, Liu AQ, Ai Y. 2015. The Poisson distribution and beyond: methods for microfluidic droplet production and single cell encapsulation. Lab Chip 15:3439–3459. https://doi.org/10.1039/c5lc00614g.
38. Loveday EK, Svinti V, Diederich S, Pasick J, Jean F. 2012. Temporal- and strain-specific host microRNA molecular signatures associated with swine-origin H1N1 and avian-origin H7N7 influenza A virus infection. J Virol 86:6109–6122. https://doi.org/10.1128/JVI.06892-11.
39. Köster S, Angilè FE, Duan H, Agresti JJ, Wintner A, Schmitz C, Rowat AC, Merten CA, Pisignano D, Griffiths AD, Weitz DA. 2008. Drop-based microfluidic devices for encapsulation of single cells. Lab Chip 8:1110–1115. https://doi.org/10.1039/b802941e.
40. Liu H, Li M, Wang Y, Piper J, Jiang L. 2020. Improving single-cell encapsulation efficiency and reliability through neutral buoyancy of suspension. Micromachines 11:94. https://doi.org/10.3390/mi11010094.
41. Gruner P, Riechers B, Semin B, Lim J, Johnston A, Short K, Baret JC. 2016. Controlling molecular transport in minimal emulsions. Nat Commun 7:10392. https://doi.org/10.1038/ncomms10392.

42. Chen Y, Wijaya Gani A, Tang SK. 2012. Characterization of sensitivity and specificity in leaky droplet-based assays. Lab Chip 12:5093–5103. https://doi.org/10.1039/c2lc40624a.
43. Etienne G, Vian A, Biocanin M, Deplancke B, Amstad E. 2018. Cross-talk between emulsion drops: how are hydrophilic reagents transported across oil phases? Lab Chip 18:3903–3912. https://doi.org/10.1039/c8lc01000e.
44. Harvey WT, Carabelli AM, Jackson B, Gupta RK, Thomson EC, Harrison EM, Ludden C, Reeve R, Rambaut A, Peacock SJ, Robertson DL, COVID-19 Genomics UK (COG-UK) Consortium. 2021. SARS-CoV-2 variants, spike mutations and immune escape. Nat Rev Microbiol 19:409–424. https://doi.org/10.1038/s41579-021-00573-0.
45. Duffy DC, McDonald JC, Schueller OJA, Whitesides GM. 1998. Rapid prototyping of microfluidic systems in poly(dimethylsiloxane). Anal Chem 70:4974–4984. https://doi.org/10.1021/ac980656z.
46. Anna SL, Bontoux N, Stone HA. 2003. Formation of dispersions using "flow focusing" in microchannels. Appl Phys Lett 82:364–366. https://doi.org/10.1063/1.1537519.
47. R Development Core Team. 2015. R: a language and environment for statistical computing (version 3.2.2). R Foundation for Statistical Computing, Vienna, Austria.
48. Bates D, Mächler M, Bolker B, Walker S. 2015. Fitting linear mixed-effects models using lme4. J Stat Soft 67:1–48. https://doi.org/10.18637/jss.v067.i01.
49. Weingartl HM, Berhane Y, Hisanaga T, Neufeld J, Kehler H, Emburry-Hyatt C, Hooper-McGreevy K, Kasloff S, Dalman B, Bystrom J, Alexandersen S, Li Y, Pasick J. 2010. Genetic and pathobiologic characterization of pandemic H1N1 2009 influenza viruses from a naturally infected swine herd. J Virol 84:2245–2256. https://doi.org/10.1128/JVI.02118-09.
50. Shu B, Wu KH, Emery S, Villanueva J, Johnson R, Guthrie E, Berman L, Warnes C, Barnes N, Klimov A, Lindstrom S. 2011. Design and performance of the CDC real-time reverse transcriptase PCR swine flu panel for detection of 2009 A (H1N1) pandemic influenza virus. J Clin Microbiol 49:2614–2619. https://doi.org/10.1128/JCM.02636-10.

CHAPTER FOUR

INFLUENZA A VIRAL BURST SIZE FROM THOUSANDS OF INFECTED SINGLE-CELLS USING DROPLET QUANTITATIVE PCR (dqPCR)

Contribution of Authors and Co-Authors

Manuscript in Chapter 4

Author: Mallory M. Thomas

Contributions: Designed studies, Performed Experiments, Analyzed data, Wrote Manuscript.

Co-Author: Geoffrey K. Zath

Contributions: Designed studies, Performed Experiments, Analyzed data, Wrote Manuscript.

Co-Author: Emma K. Loveday

Contributions: Designed studies, Wrote Manuscript.

Co-Author: Dimitri A. Bikos

Contributions: Performed Experiments.

Co-Author: Steven Sanche

Contributions: Analyzed data

Co-Author: Ruian Ke

Contributions: PI on the Study, Acquired funding, Designed studies, Analyzed data.

Co-Author: Christopher B. Brooke

Contributions: PI on the Study, Acquired funding, Designed studies, Provided Virus.

Co-Author: Connie B. Chang

Contributions: PI on the Study, Acquired funding, Designed studies, Wrote Manuscript.

Manuscript Information

Mallory M. Thomas, Geoffrey K. Zath, Emma K. Loveday, Dimitri A. Bikos, Steven Sanche, Ruian Ke, Christopher B. Brooke, Connie B. Chang.

PLOS Pathogens

Status of Manuscript:
☐ Prepared for submission to a peer-reviewed journal
☐ Officially submitted to a peer-reviewed journal
☐ Accepted by a peer-reviewed journal
☒ Published in a peer-reviewed journal

Public Library of Science (PLOS)

July 1, 2024

DOI 10.1371/journal.ppat.1012257

PLOS PATHOGENS

RESEARCH ARTICLE

Influenza A viral burst size from thousands of infected single cells using droplet quantitative PCR (dqPCR)

Geoffrey K. Zath[1,2☯], Mallory M. Thomas[1,3,4☯], Emma K. Loveday[1,2], Dimitri A. Bikos[1,2], Steven Sanche[5], Ruian Ke[5], Christopher B. Brooke[6], Connie B. Chang[3]*

1 Center for Biofilm Engineering, Montana State University, Bozeman, Montana, United States of America, **2** Department of Chemical and Biological Engineering, Montana State University, Bozeman, Montana, United States of America, **3** Department of Physiology and Biomedical Engineering, Mayo Clinic, Rochester, Minnesota, United States of America, **4** Department of Microbiology and Cell Biology, Montana State University, Bozeman, Montana, United States of America, **5** Theoretical Biology and Biophysics, Los Alamos National Laboratory, Los Alamos, New Mexico, United States of America, **6** Department of Microbiology, University of Illinois, Urbana-Champaign, Urbana, Illinois, United States of America

☯ These authors contributed equally to this work.
* chang.connie@mayo.edu

Citation: Zath GK, Thomas MM, Loveday EK, Bikos DA, Sanche S, Ke R, et al. (2024) Influenza A viral burst size from thousands of infected single cells using droplet quantitative PCR (dqPCR). PLoS Pathog 20(7): e1012257. https://doi.org/10.1371/journal.ppat.1012257

Editor: Anice C. Lowen, Emory University School of Medicine, UNITED STATES

Received: August 22, 2023

Accepted: May 13, 2024

Published: July 1, 2024

Data Availability Statement: Data supporting this study are included within the article and/or supporting materials. The datasets generated during and/or analyzed during the current study are available on the Chang Lab GitHub repository at https://github.com/chang-lab-mayo-clinic/Influenza-A-Burst-Size.

Funding: This work was supported by Defense Advanced Research Projects Agency (DARPA) grant W911NF-17-2-0034 to CBC, RK and CBB (https://www.darpa.mil/); by the National Institutes

Abstract

An important aspect of how viruses spread and infect is the viral burst size, or the number of new viruses produced by each infected cell. Surprisingly, this value remains poorly characterized for influenza A virus (IAV), commonly known as the flu. In this study, we screened tens of thousands of cells using a microfluidic method called droplet quantitative PCR (dqPCR). The high-throughput capability of dqPCR enabled the measurement of a large population of infected cells producing progeny virus. By measuring the fully assembled and successfully released viruses from these infected cells, we discover that the viral burst sizes for both the seasonal H3N2 and the 2009 pandemic H1N1 strains vary significantly, with H3N2 ranging from 10^1 to 10^4 viruses per cell, and H1N1 ranging from 10^1 to 10^3 viruses per cell. Some infected cells produce average numbers of new viruses, while others generate extensive number of viruses. In fact, we find that only 10% of the single-cell infections are responsible for creating a significant portion of all the viruses. This small fraction produced approximately 60% of new viruses for H3N2 and 40% for H1N1. On average, each infected cell of the H3N2 flu strain produced 709 new viruses, whereas for H1N1, each infected cell produced 358 viruses. This novel method reveals insights into the flu virus and can lead to improved strategies for managing and preventing the spread of viruses.

Author summary

Viruses infect and exploit host cells to reproduce and spread. The viral burst size, or the number of viral particles released from an infected cell, plays a critical role in understanding infection dynamics and overall viral fitness. However, accurately determining burst size for many single cells using conventional laboratory methods can be challenging.

of Health (NIH) National Institute of Allergy and Infectious Diseases (NIAID) grant 1R56AI156137-01 to CBC (https://www.niaid.nih.gov/); and by the National Science Foundation (NSF) CAREER grant 1753352 to CBC (https://www.nsf.gov/). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Here, we introduce dqPCR, a droplet microfluidic method for the rapid measurement of influenza virus numbers produced by thousands of individual cells. Our findings revealed that only a small proportion of infected cells are responsible for producing a significant portion of the total viral population. By utilizing this method in future studies, we can gain a deeper understanding of the role of diversity in rapidly evolving viruses.

Introduction

The number of viral particles produced from an infected host cell is known as the viral burst size. First characterized in bacteriophage populations [1,2], burst size is an important measure of viral fitness and plays an important role in modeling of viral infection dynamics [3]. To study viral infections, tissue culture systems or animal models are typically used; however, these methods provide an average measurement of infected cells and may not capture the variation between individual infected cells [4]. In the case of influenza A virus (IAV) infection, individual cells exhibit significant heterogeneity in production of infectious virions [5,6], viral gene expression [6–9], and host innate immune activation [8]. Previous estimates of IAV burst size indicate a similar trend; average IAV burst size measured from bulk tissue culture was estimated at 11.5 plaque forming units (pfu) per cell [10] while single cell measurements ranged from 1 to 1000 pfu/cell [5]. However, these estimates relied on plaque assays and were limited to quantification of fully infectious viral particles. The plaque assays overlooked the majority of non-infectious particles that make up 90–99% of virions in the IAV population that rely on co-infection to propagate [11,12]. Therefore, new techniques are needed to accurately determine the IAV burst size by capturing the full range of particle types released from single-cell infections.

Various methods have been used to measure IAV production from single-cell infections, including RT-PCR [5,6,13], RNA sequencing [6–9], flow cytometry [11], and infectivity titer [5,6,13]. However, each method has limitations in accurately measuring viral burst size. Infectivity assays only detect fully infectious viral particles and cannot account for incomplete viral particles [11,12] that contribute to viral replication and host immune responses [14]. RT-PCR, RNA sequencing, and flow cytometry typically measure intracellular viral replication, neglecting the packaging and release of extracellular viral particles. These methods are limited because intracellular viral transcription may correlate poorly with viral progeny production [9]. Additionally, these methods can be low-throughput, capturing only a limited number of single-cell infections [5–7,9,13], and may not be sufficient to study highly diverse viral populations like IAV. Determining IAV burst size distributions, therefore, requires a high-throughput method capable of measuring extracellular viral particles released during single-cell infection. One such method is drop-based microfluidics. This flexible platform enables high-throughput, single cell virology experimentation in isolated microenvironments [15–20]. Recently, we demonstrated that IAV can replicate in single cells encapsulated in drops and is comparable to bulk infection [21].

Existing drop-based microfluidic methods for quantifying genome copies rely on endpoint measurements like droplet digital PCR (ddPCR) or real-time fluorescence measurements using qPCR on a microfluidic chip. However, these approaches have limitations. ddPCR can only confirm the presence or absence of nucleic acids [22,23] and cannot quantify the contents of an individual drop. qPCR on a chip can quantify the contents of an individual drop, but requires complex microfluidic devices and results are relatively low-throughput [24–30]. Improving the throughput of drop-based qPCR would allow for rapid measurement of a significantly larger number of drops, thus enabling the study of large and diverse viral populations.

In this study, we introduce a microfluidic method called droplet quantitative PCR (dqPCR), which combines our prior work [21,31] with flow-based droplet fluorescence detection. The dqPCR method is used to measure extracellular burst size distributions of thousands of IAV-infected single cells that were encapsulated into drops. Human alveolar epithelial (A549) cells were infected with either influenza A/Perth/16/2009 (H3N2; human seasonal isolate) or influenza A/California/07/2009 (H1N1; 2009 pandemic strain) at multiplicity of infection (MOI) of 0.1. The IAV burst sizes were found to be highly diverse, spanning three orders of magnitude (10^1 to 10^4 viruses/cell) for H3N2 and two orders of magnitude (10^1 to 10^3 viruses/cell) for H1N1 across thousands of single cell infection measurements. IAV production followed a negative-binomial distribution, with modeled mean burst sizes of 709 and 358 viruses/drop for H3N2 and H1N1, respectively. Notably, we found that 10% of cells produced ≈60% or ≈40% of M gene RNA during H3N2 or H1N1 infections, respectively. The dqPCR method provides a novel high-throughput approach for quantifying released viruses from infected single cells and introduces an unprecedented means for characterizing the infection dynamics of highly heterogeneous viral populations.

Results

Drop-based microfluidic platform for measuring viral burst size from single-cell infections

To determine IAV burst size from single cell infections, H3N2 and H1N1 strains were first inoculated onto A549 cells in bulk culture with an MOI of 0.1 pfu/cell (Fig 1A). A low MOI was chosen to ensure that of the few cells that were infected, most cells were infected with a single virion, as variation in cellular MOI could affect our measurements due to co-infection [32]. While a low MOI of 0.1 is expected to result in only a small portion of the infected cells producing progeny virus, the high-throughput capability of dqPCR enables the screening of tens of thousands of drops. Single infected cells were then encapsulated into 100 μm diameter microfluidic drops, such that most drops were either empty or contained a single cell (Fig 1B). Drops were incubated for 18 hours post infection (hpi) to allow for replication of IAV progeny virus (Fig 1C) with numbers comparable to bulk infection (S1 Fig). Interestingly, the supernatant collected from drops produced slightly more progeny virus than bulk (2.4× to 6.8× greater production, $p << 0.05$). The burst sizes of single-cell infections were quantified by the number of IAV genome copies (cpd), which can be equated to viruses per cell for each drop that contains an infected cell producing progeny virus. We note that these infected cells that produced progeny virus represent a small portion of the total infected cell population. The genome copies were measured using a multiplexed RT-qPCR assay targeting two template sequences, the IAV M gene and cellular β-actin (S1 Table), following our previous work [31]. A microfluidic chip [20] was used to split off a $1/8^{th}$ portion of the drop containing progeny virus and merge with a continuous flow of RT-qPCR mix (Fig 1D). This sampling method isolates extracellular viral progeny released from the infected host cell, allowing us to quantify genome copies only from fully assembled and released extracellular virions. Note that our recent work indicates minimal amplification efficiency loss when directly amplifying infected supernatant in drops compared to bulk qPCR [21]. Based on this finding, we assume that using an M gene stock serves as a suitable standard for comparing our infected drop samples. Drops containing progeny virus and RT-qPCR mix were thermocycled using a standard qPCR machine and removed at successive PCR cycle numbers, N (Fig 1E). Normalized drop fluorescence intensities (ΔR_N, S5 Materials and Methods) were measured by flowing drops through a microfluidic device mounted on a custom optical detection microscope [33] (Fig 1F). This flow-based detection method significantly increased the throughput of our drop fluorescence measurements during

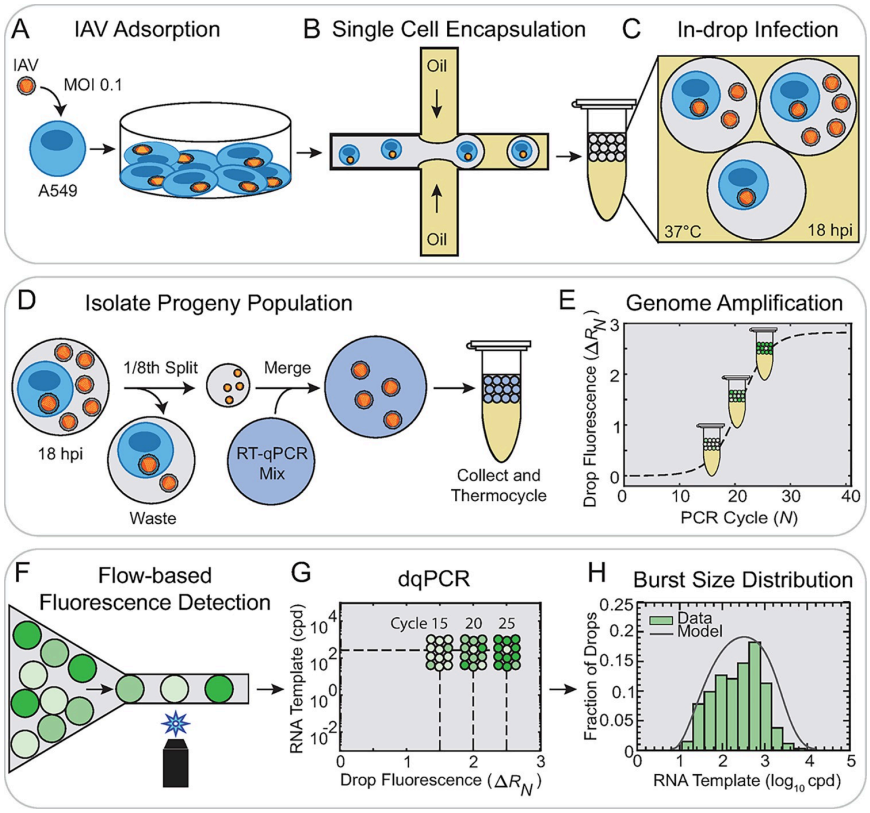


Fig 1. Drop-based Microfluidic Platform for Measuring Viral Burst Size from Single-Cell Infections. (A) IAV (H1N1 or H3N2) was inoculated onto plated A549 cells at an MOI of 0.1 pfu/cell. **(B)** Single infected cells were encapsulated into 100 μm drops using a flow-focusing device. **(C)** Drops were incubated at 37˚C for 18 hpi to allow for replication of IAV progeny virus. **(D)** In-drop infections were injected into a microfluidic device that splits 1/8th of the extracellular progeny virus from the host cell and merges these viruses with a continuous flow of multiplexed RT-qPCR mix targeting IAV M gene RNA and cellular β-actin mRNA. **(E)** Drops were thermocycled in a standard qPCR machine and sampled at discontinuous *N*. **(F)** ΔR_N was measured serially using flow-based detection. **(G)** ΔR_N values were converted to RNA cpd using dqPCR. **(H)** IAV burst size distributions of M gene RNA cpd from thousands of single-cell infections (green bars). The mean burst size of each distribution was calculated from a parametric model (gray curve) fitted to our data.

https://doi.org/10.1371/journal.ppat.1012257.g001

burst size experiments compared to our previous static image analyses [31]. ΔR_N of filtered drops were converted to M gene and β-actin cpd using dqPCR (Fig 1G). A variable filter was applied to drops containing high amounts of β-actin to exclude them from the final burst size distributions (S8 Results). These drops were assumed to be contaminated with intracellular lysate and were not included in further analysis. The abundance of M gene RNA across thousands of single-cell infections revealed a wide distribution of IAV burst sizes (Fig 1H).

Droplet Quantitative PCR (dqPCR) method for converting drop fluorescence to nucleic acid concentration

The dqPCR method allows for the construction of standard curves that directly relate ΔR_N (normalized drop fluorescence) to M gene RNA copy number (cpd). In traditional RT-qPCR, standard curves plot the cycle threshold (*Ct*) at which reporter fluorescence rises above

background for a dilution series of known RNA template concentrations (S7 Materials and Methods). This method requires continuous measurements of fluorescence amplification across all *N*. However, in dqPCR, drops are sampled discontinuously and only require evaluation of sample fluorescence at a single *N*.

To construct standard curves at a single *N* using dqPCR, we generated a reference amplification curve for a known RNA template concentration amplified in drops (Fig 2A, solid blue curve, 1.71×10^2 M gene cpd). Drop ΔR_N was measured at $N = 1$, 10, 13, 16, 19, 22, 28 and 40 (Fig 2A, black dots). A continuous amplification curve was fit to the discontinuous drop fluorescence measurements using sigmoidal curve fitting using PCR efficiency (SCF-E) (S8 Materials and Methods). The model determined four PCR efficiency (E_N) parameters that best fit

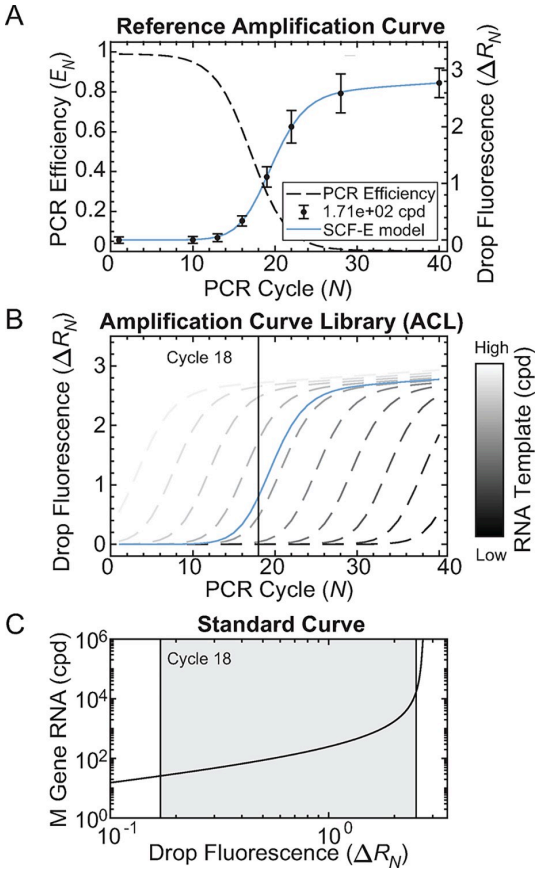


Fig 2. Droplet Quantitative PCR (dqPCR) Model for Converting Drop Fluorescence to Nucleic Acid Concentration. (A) M gene RNA (1.71×10^2 cpd) was amplified in 50 μm drops. ΔR_N was detected at $N = 1$, 10, 13, 16, 19, 22, 28, 40 and displayed as the mean (black dots) with error bars representing one standard deviation. Approximately 1000 drops were detected at each cycle. A continuous reference amplification curve (blue solid curve) was reconstructed from discontinuous ΔR_N measurements using SCF-E. The E_N curve (black dashed curve) was fit to ΔR_N at sampled cycle numbers to calculate ΔR_N at non-sampled cycle numbers. **(B)** The efficiency parameters fit from the reference amplification curve (blue solid curve) were used to model 1000 virtual amplification curves (gray dashed curves, subset of 10 shown), creating a high-resolution dilution series called an Amplification Curve Library (ACL), representing 1000 unique M gene RNA concentrations. **(C)** Standard curves were constructed using ΔR_N from all virtual curves at a single *N* (*i.e.*, cycle 18). The valid range of the standard curve (gray shaded area) is determined by the min and max ΔR_N of the reference curve.

https://doi.org/10.1371/journal.ppat.1012257.g002

the ΔR_N values of the sampled N (E_{max}, E_{min}, $N_{0.5}$, and k). Once fit, these parameters were used to calculate the E_N curve for all N (Fig 2A, dashed black curve). The E_N curve was then used to produce a ΔR_N reference amplification curve that included non-sampled N (Fig 2A, solid blue curve). SCF-E was validated for bulk RT-qPCR of M gene RNA (S5A Fig) across five orders of magnitude (S2 Results and S2 Table). Additionally, SCF-E was validated for drops of 1.71×10^2 M gene cpd (S12A Fig).

The E_N parameters obtained from the reference amplification curve were used to model virtual amplification curves, creating a high-resolution dilution series referred to as an Amplification Curve Library (ACL) (S9 Materials and Methods). The ACL simulated 1000 virtual amplification curves, each with a unique RNA template concentration (Fig 2B, dashed gray curves, subset of 10 shown). Once constructed, the ACL was used to generate a standard curve that related RNA template concentration to ΔR_N at a single N (Fig 2C). This was done by recording ΔR_N of every virtual curve at a single N (*e.g.*, $N = 18$) where unknown drops were sampled (Fig 2B, black vertical line). The valid range for conversion with the standard curve (Fig 2C, gray shaded area) was determined by the ΔR_N values of the reference curve (Fig 2A). The maximum valid ΔR_N was calculated as the mean minus the standard deviation of ΔR_N at $N = 40$, which removes drops where amplification has plateaued. The minimum valid ΔR_N was set at the 99th percentile of ΔR_N at $N = 1$ to remove background fluorescence. ACL standard curves were shown to be robust across N for both bulk (S4 Results) and drops (S12B Fig). Additionally, we tested ACL standard curves for human and plant DNA template sequences (S5 Results) and found that ΔR_N accurately converted to genome concentrations within a two-fold change of the expected means. Together, the SCF-E model for generating reference amplification curves and the ACL method of constructing standard curves form the dqPCR analysis pipeline that enables the conversion of drop fluorescence to cpd (Fig 2).

Resolving IAV M Gene RNA concentrations from single and mixed droplet populations

To validate the capability of dqPCR to resolve a range of RNA concentrations in drops, we conducted a control experiment amplifying three known concentrations of M gene RNA (1.71×10^1, 1.71×10^2, and 1.71×10^3 cpd) in 50 μm drops (Fig 3A). These RNA concentrations were selected because they fall within the expected range of IAV single-cell burst sizes from previous estimates [5,9,10]. ΔR_N was measured for each group individually at multiple intermediate N and converted to M gene cpd using dqPCR (S6 Table). A new reference amplification curve of 1.71×10^2 M gene cpd was generated for each M gene group, as PCR kinetics can vary from day to day (S13 Fig). The measured mean M gene cpd at each sampled N for all concentrations fell within a 2-fold change of the expected concentration (S14 Fig). A linear mixed-effects model showed minimal deviation from the expected concentration at different cycle numbers (<1% difference/cycle, S6 Results).

Thus, M gene cpd measurements from all N were pooled together into single distributions with means of 2.03×10^1 cpd (red, $n = 4{,}513$ drops), 1.31×10^2 (blue, $n = 15{,}899$ drops), 1.16×10^3 (green, $n = 65{,}249$ drops) (Fig 3B). Distributions were centered around the expected mean and exhibited a linear relationship between measured and expected cpd, falling within a 2-fold change of the expected mean (S17 Fig). The heterogeneity of each distribution was examined using the Gini coefficient (G) (Fig 3C); a value typically used to measure the degree of wealth inequality and ranges from 0 to 1 [34]. A G value of 0 indicates total equality, where each cell produces the same number of viruses, while a G value of 1 signifies absolute inequality, where one cell produces all the viruses. Thus, a higher G corresponds to a more heterogeneous distribution. The uniformity of each M gene distribution (Fig 3C) was as follows: 10^1

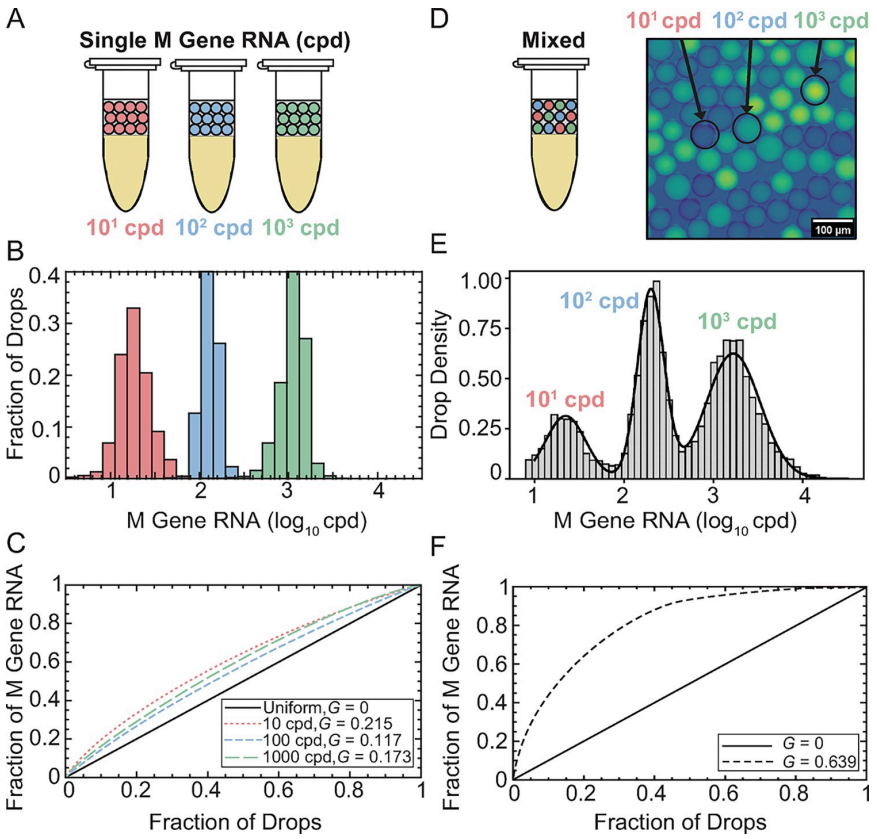


Fig 3. Resolving IAV M Gene RNA Concentrations from Single and Mixed Droplet Populations. (A) Three known concentrations of M gene RNA, 1.71×10^1 cpd (red), 1.71×10^2 cpd (blue), and 1.71×10^3 cpd (green) were amplified in 50 μm drops. **(B)** ΔR_N was detected at multiple different N (S6 Table) within each group and converted to M gene cpd. M gene cpd measurements were pooled to a single distribution for each group, with means of 2.03×10^1 cpd (red, $n = 4{,}513$ drops); 1.31×10^2 (blue, $n = 15{,}899$ drops); 1.16×10^3 (green, $n = 65{,}249$ drops). **(C)** The uniformity of each distribution was quantified using G. **(D)** Three M gene RNA concentrations in drops were collected in a mixed sample. **(E)** Mixed ΔR_N were detected at $N = 16$–19, converted to M gene cpd, and pooled to a single distribution (grey bars, $n = 68{,}748$). Three individual peaks were isolated from the mixed distribution using GMM analysis (black curve) (S7 Table). **(F)** The heterogeneity of the mixed population compared to the individual populations in (C) was significantly greater with $G = 0.639$.

https://doi.org/10.1371/journal.ppat.1012257.g003

cpd ($G = 0.215$), 10^2 cpd ($G = 0.117$) or 10^3 cpd ($G = 0.173$). These values represent the baseline level of noise for single concentrations M gene RNA in drops ($\sim G < 0.22$) (Fig 3C).

To further confirm the ability of dqPCR to resolve a mixed drop population, we conducted another control experiment in which the same three known concentrations of M gene RNA (1.71×10^1, 1.71×10^2, or 1.71×10^3 cpd) were separately encapsulated into 50 μm drops and collected as a mixed sample prior to thermocycling (Fig 3D). ΔR_N of the mixed drops was detected at $N = 16$–19 and converted to M gene cpd, enabling high-throughput analysis with over 69,000 drops sampled (Fig 3E). The data from the mixed sample yielded three individual distributions and were isolated using a Gaussian mixture model (GMM) (S7 Results and S7 Table). G for the mixed drop sample containing 10^1 to 10^3 M gene cpd was 0.639 (Fig 3F), demonstrating greater heterogeneity compared to the measurement of single RNA concentrations in drops, as expected. These results demonstrate the ability to resolve three M gene RNA concentrations from a mixed concentration of drops.

H3N2 and H1N1 single-cell infection burst size distributions

To determine the burst size distributions of H3N2 and H1N1 populations, we measured M gene RNA abundance in $>10^4$ drops containing the small portion of infected cells that successfully produced progeny virus. Infected drops were processed by isolating $1/8^{th}$ of a drop containing progeny viruses and merging it with RT-qPCR mix. The merged drops were thermocycled and analyzed using flow-based fluorescence detection. To ensure that the experimental burst size measurements only included extracellular viral progeny and did not include intracellular viral RNA from cell lysate, a multiplexed RT-qPCR assay was used to detect the presence of an abundant cellular protein, β-actin (S18 Fig). Drops with high levels of β-actin were filtered out from the final burst size distributions using a variable filter (S8 Results). Filtered drops were quantified at N = 16, 19, 22, 25, and 28 (S8 Table) using dqPCR and pooled to create single burst size distributions. For each biological replicate, a new SCF-E reference curve (S20 Fig) was used for dqPCR (S9 Table). As an additional control to ensure that the measurements included only extracellular viral progeny, single IAV-infected cells in drops were treated with oseltamivir, a drug preventing the cleavage of assembled viral particles from the host membrane. Infected cells treated with oseltamivir exhibited M gene RNA abundance comparable to mock infected cells at 18 hpi. (S9 Results, S21 Fig).

Experimental IAV burst sizes revealed increased progeny virus production in the H3N2 population compared to H1N1, as observed from a random sample of 100 drops and plotting in ascending order of M gene RNA (Fig 4A). The burst size distributions were highly diverse, ranging from ~10^1 to ~10^3 M gene cpd for H1N1 (n = 5,585 drops) and an order of magnitude higher, from ~10^1 to ~10^4 M gene cpd, for H3N2 (n = 9,620 drops) (Fig 4B). The mean of the H3N2 distribution was 1.0×10^3 M gene cpd (Fig 4B, blue bars). This fell outside of the interquartile range (IQR) of the distribution, with lower (25^{th} percentile) and upper (75^{th} percentile) bounds of 93 and 890 M gene cpd, respectively (S8 Table). The mean can fall outside of

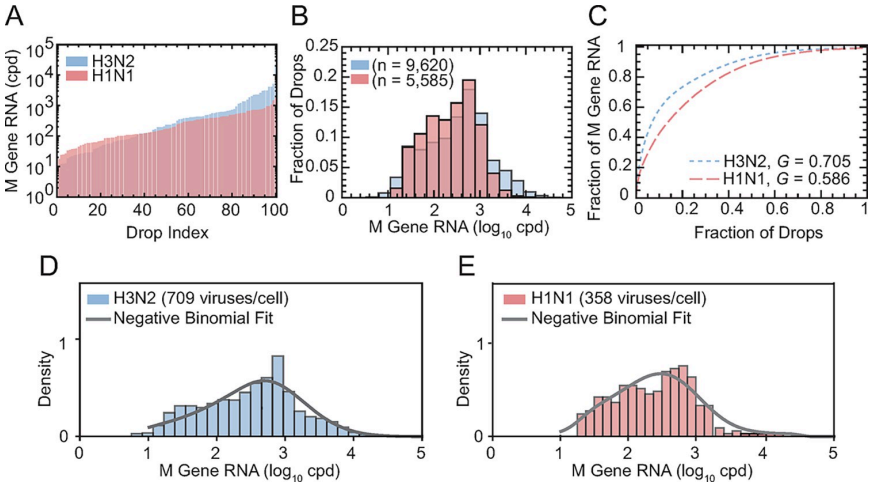


Fig 4. H3N2 and H1N1 Single-Cell Infection Burst Size Distributions. IAV burst sizes of H3N2 (blue) and H1N1 (red) were sampled at N = 16, 19, 22, 25, and 28 across three biological replicates (S8 Table). **(A)** Low to high M gene RNA (cpd) from a random sample of drops (n = 100) for H3N2 (blue) and H1N1 (red). **(B)** Single burst size distributions for H3N2 (blue, n = 9,620) and H1N1 (red, n = 5,585) populations. The measured mean burst sizes were 1.0×10^3 M gene cpd (H3N2) and 4.8×10^2 M gene cpd (H1N1). **(C)** The H3N2 distribution (G = 0.705) was less uniform than H1N1 (G = 0.586). **(D-E)** Distributions were fitted to a negative binomial (grey curve, S10 Table). The modeled mean burst sizes were (D) 709 cpd (H3N2) and (E) 385 cpd (H1N1).

https://doi.org/10.1371/journal.ppat.1012257.g004

the IQR in a highly skewed distribution, which occurs when a small fraction of the infected cell population produces a greater proportion of virus. The mean of the H1N1 burst size distribution was 4.8×10^2 M gene cpd (Fig 4B, red bars), falling within the IQR of 88 and 630 M gene cpd (S8 Table). This suggests less heterogeneity in the H1N1 population, which is further supported by examining the uniformity of each distribution using *G* (Fig 4C). The H3N2 distribution is less uniform ($G_{H3N2} = 0.705$) compared to H1N1 ($G_{H1N1} = 0.586$). In both viruses, a minor population of infected cells produces a large fraction of progeny virus. For H3N2, 10% of infected cells produced ≈60% of measured progeny virus (Fig 4C, blue dotted line). For H1N1, 10% of infected cells produced ≈40% of progeny virus (Fig 4C, red dashed line).

To estimate the most accurate mean burst size, we apply a statistical model (S10 Results) that considers both measurement noise and distribution shape. The model demonstrated that the burst size data is best fit by a negative-binomial distribution (Fig 4D–4E). The shape parameter of the negative binomial distribution was estimated to be 0.48 and 0.49 for H3N2 and H1N1, respectively (S10 Table). This indicates extreme cell-to-cell variability in burst sizes for both H3N2 and H1N1 populations, as the negative-binomial distribution is considered highly dispersed when the shape parameter is less than 1. Using the negative binomial fit, we estimated that the modeled mean burst sizes for each population were 709 viruses/cell for H3N2 and 358 viruses/cell for H1N1 (Fig 4D–4E).

Discussion

We introduce a droplet microfluidic method called dqPCR to quantify viral burst size from thousands of single infected cells that successfully produced progeny virus. dqPCR is a quantitative PCR technique for droplets that does not require specialized microfluidic devices or continuous measurements of a sample. This has led to a drastic improvement in throughput compared to previous on-chip qPCR work [24–30]. Compared to ddPCR, which measures only single genome copies per drop, dqPCR can resolve a range of genome copies from a single copy to up to thousands of copies per drop.

Estimating viral burst sizes using population averaging masks cell-to-cell variability. Additionally, single-cell burst size estimates based upon infectivity, like the plaque assay, do not account for the production of incomplete particles that are abundant in IAV populations [11,12]. Here, we account for both infectious and non-infectious particles released from single infected cells by quantifying viral genome abundance. This method does not measure particle number from all infected cells in the population, but rather only those infected cells that successfully produced and released progeny virus into the extracellular environment. Future work could combine simultaneous tracking of cellular infection with burst size quantification to reveal the true proportion of the infected cell population that produce IAV progeny particles. Nonetheless, we account for all single cells in the productive infection population, without averaging, to investigate whether one cell produces more progeny or if all cells produce equal progeny. We measured IAV burst sizes across $>10^4$ drops containing cells producing H3N2 or H1N1 progeny virus. This has yielded the largest estimate of single-cell viral burst size to date. The range of our burst size measurements spanned from 10^1 to 10^4 viruses/drop, which is an order of magnitude wider than previous, lower throughput ($n = 430$) estimates of H1N1 infectious virus production in single cells (10^1 to 10^3 pfu/cell) [5]. This is likely due to our method that measures all viral genome copies rather than infectious particles. A mean burst size of 709 and 358 viruses/cell was modeled for H3N2 and H1N1 strains, respectively, from highly dispersed negative-binomial distributions. The mean burst size for H3N2 (709 viruses/cell) compares similarly with a previous mean estimate of 962 viruses/cell for H3N2 in a bulk culture [10]. One caveat of this study is that some drops contained more than one cell per drop. Across

both populations, for a drop that contained a cell or cells, an average of 86% of these drops were loaded with 1 cell, 12% with 2 cells, and 2% with 3 cells or more (S23 Fig). Another limitation of this study is that the cells were not synchronized to the same growth stage, which may influence viral numbers. Future work can improve upon our methods to encapsulate exactly one cell per drop [35,36] and explore cell synchronization.

The observed heterogeneity in the single cell burst sizes (G_{H3N2} = 0.705 and G_{H1N1} = 0.586) indicates that a minor population (10%) of infected host cells produced a large fraction of progeny virus (≈60% for H3N2 and ≈40% for H1N1) (Fig 4C). While our burst size measurements quantify the number of M gene segments, similar heterogeneity has been observed when quantifying the number of HA and NA segments using single-cell sequencing (G = 0.78), where 6.5% of infected cells generated over 50% of the counted progeny virus [9], although this study examined fewer single cells (n = 92). Additionally, these trends align with single-cell transcriptomic studies of IAV infection, which show that 10–15% of infected cells are responsible for 50% of viral gene expression [7], although levels of intracellular transcripts are not well correlated to progeny virus production [9].

This study is the first to compare viral burst sizes of two different clinically relevant human IAV strains and provides the first burst size estimate for the H3N2 strain, revealing differences in progeny production between the two strains. Future research could investigate specific viral and host factors that contribute to the observed heterogeneity in viral output, both within and between IAV strains, as well as the extent to which *in vitro* measurements reflect burst sizes across different host cell types *in vivo*. As dqPCR is a multiplexed assay, it can be adapted to simultaneously measure multiple template sequences, such as other genome segments within IAV or both virus and host templates. In the case of IAV, dqPCR could be applied to investigate segment reassortment [37] or the frequency of incomplete viral genomes [10–12], or it could be extended to study genome recombination in other viruses [18]. The dqPCR method presented here not only expands our current toolset for high-throughput single-cell characterization, but also enables the detection of nucleic acid abundance on an unprecedented scale. Future research can leverage the highly resolved single-cell data sets provided by dqPCR to further characterize host-virus interactions, improve diagnostics, or develop more effective vaccines and anti-viral treatments.

Materials and methods

Virus strains and cell lines

Influenza A virus (IAV) strains H1N1 (A/California/07/09) and H3N2 (A/Perth/16/2009) were obtained as seed stock from Dr. Christopher Brooke (University of Illinois Urbana-Champaign) and expanded in Madin-Darby canine kidney cells (MDCK) and human alveolar epithelial cells (A549) obtained from ATCC (CCL-34, CCL-185). MDCK cells were maintained in Dulbecco's Modified Eagle medium (Corning) supplemented with 10% fetal bovine serum (FBS) (HyClone) and 1× penicillin-streptomycin (pen/strep) (Fisher Scientific). A549 cells were maintained in Hams F-12 medium (Corning), also supplemented with 10% FBS and 1× pen/strep. IAV seed stocks were generated using the JetPRIME kit (Polyplus) to co-transfect Human kidney epithelial-like cells (293T) with 8 plasmids each carrying the sequence of one IAV genome segment: pDZ::PB2, pDZ::PB1, pDZ::PA, pDZ::HA, pDZ::NP, pDZ::NA, pDZ::M, and pDZ::NS. Plaque isolates from transfection supernatants were then amplified in MDCK cells to produce IAV seed stocks. The seed stock virus population (P1) was expanded into working stocks (P2) for both the H1N1 and H3N2 strains. For expansion, an MOI of 0.0001 was used with an expansion time of 3 dpi based on cytopathic effect. The working stocks were used for all experiments and generated by infecting MDCK cells with seed stock at

an MOI of 0.01 pfu/cell. Working stock titers were determined by plaque assay on MDCK cells. IAV burst sizes were measured from single-cell infection of A549 cells at an MOI of 0.1 pfu/cell.

Microfluidic device fabrication

Microfluidic devices were fabricated using polydimethylsiloxane (PDMS) (Sylgard 184) by patterning SU-8 photoresist on silicon wafers using standard soft photolithography.

IAV infection

A549 cells were seeded onto 6-well plates at a concentration of 1×10^6 cells/well and incubated at 37˚C with 5% CO_2 for 24 hrs. Enough wells were plated to collect a final cell suspension of 2×10^6 cells/mL following infection. For our experiments, we plated cells onto 6 wells for infection samples and 4 wells for mock infection controls. Confluent cell monolayers were washed twice with 1× PBS (Corning) and infected at an MOI of 0.1 with H1N1 or H3N2 working stocks suspended in A549 infection media (Ham's F-12 medium supplemented with 1 mM HEPES (HyClone), 1× pen/strep, 0.2% Bovine Serum Albumin (BSA) (MP Biomedical) and 1 μg/mL TPCK-trypsin (Worthington Biomedical). Infected cells were incubated at 37˚C for 1 hr. Mock infected cells were treated with media without the virus and otherwise processed the same as infected cells. After 1 hr, supernatants were removed and replaced with 2 mL of fresh media for an additional 1 hr incubation period at 37˚C. Following the second incubation period, cells were washed twice with 1× PBS. Cells were then detached from wells using 300 μL of 0.25% Trypsin containing 1× EDTA (Fisher Scientific) for 5 min. Next, 700 μl of A549 culture media (Hams F-12 medium supplemented with 10% FBS and 1× pen/strep) was added to the cells to neutralize trypsinization. Cells were collected into two final suspensions, infected and mock, and the wells were washed one more time with 500 μl of A549 culture media. Cell suspensions were centrifuged at 500 × *g* for 5 min to pellet, washed with 1× PBS, and centrifuged a second time. The washed cell pellets were each resuspended in infection media (Ham's F-12 media supplemented with 1 mM HEPES, 1× pen/strep and 0.2% BSA) at a final concentration of 2×10^6 cells/mL. For each IAV strain (H1N1 or H3N2), we sampled drop and bulk infections, with mock infection controls at the following timepoints: drop infection (0 and 24 hpi), bulk infection (0 and 24 hpi), mock drop infection (24 hpi), mock bulk infection (24 hpi). This required final cell suspension volumes of 3 mL for infected cells (2 mL for drop infections, 1 mL for bulk infections) and 2 mL for mock-infected cells (1.5 mL for mock drop infections, and 500 μl for mock bulk infection). For bulk infection and mock samples, 500 μl of cell suspensions were re-plated onto 3 wells for a final concentration of 1×10^6 cells/well. Re-plated cells were either incubated at 37˚C with 5% CO_2 for 24 hpi or immediately frozen at -80˚C for 0 hpi. Drop infection protocols are described in the following section.

Droplet encapsulation of infected cells

Single A549 cells infected with either H1N1 or H3N2 were encapsulated in 100 μm diameter drops using a flow-focusing drop-making device. Mock-infected cells were encapsulated using the same procedure. A 2 mL suspension of 2×10^6 cells/mL in media was loaded into a 3 mL syringe and injected into the device at 1000 μL/hr using syringe pumps (New Era NE-1000). Simultaneously, 4.5 mL of 1.5 wt% (w/w) solution of PEG-$PFPE_2$-based surfactant (RAN Biotechnologies, #008-FluoroSurfactant) in fluorinated HFE 7500 oil (3M) was loaded into a 5 mL syringe and injected at 3000 μL/hr. Drops were collected for 30 min into a 5 mL syringe. Syringes were capped and incubated at 37˚C and 5% CO_2 for 24 hpi to allow for viral replication in drops. Drop infections sampled at 0 hpi were immediately frozen overnight at -80˚C.

Droplet Quantitative PCR (dqPCR) of isolated progeny viruses

A multiplexed RT-qPCR assay was used to simultaneously amplify both IAV M gene RNA and cellular β-actin mRNA [31]. Drops containing infected cells were split and merged with RT-qPCR mix. Drops were thermocycled and drop fluorescence was measured using epifluorescence or flow-based fluorescence detection.

More detailed methods are in S1–S9 Materials and Methods files.

Supporting information

S1 Materials and Methods. RT-qPCR Template and Primer Sequences.
(PDF)

S2 Materials and Methods. Microfluidic Device Fabrication.
(PDF)

S3 Materials and Methods. Bulk RT-qPCR.
(PDF)

S4 Materials and Methods. Droplet Quantitative PCR (dqPCR) of Isolated Progeny Viruses.
(PDF)

S5 Materials and Methods. Drop Fluorescence Imaging.
(PDF)

S6 Materials and Methods. Drop Flow-based Fluorescence Detection.
(PDF)

S7 Materials and Methods. Cycle Threshold (Ct) Method of Creating Standard Curves for RT-qPCR.
(PDF)

S8 Materials and Methods. Sigmoidal Curve Fitting using PCR Efficiency (SCF-E).
(PDF)

S9 Materials and Methods. Amplification Curve Library (ACL) Method of Constructing Standard Curves for RT-qPCR.
(PDF)

S1 Results. Comparing M Gene Abundance During Bulk and Drop Infections.
(PDF)

S2 Results. Validating SCF-E with Bulk RT-qPCR of IAV M Gene.
(PDF)

S3 Results. Setting a PCR Efficiency Constant (E_N) for the ACL.
(PDF)

S4 Results. Validating ACL Standard Curves for Multiple PCR Cycle Numbers.
(PDF)

S5 Results. Validating ACL Standard Curves for Multiple Template Sequences.
(PDF)

S6 Results. Validating dqPCR for Multiple PCR Cycle Numbers.
(PDF)

S7 Results. Extracting Individual Distributions from Mixed Droplet Detection Data.
(PDF)

S8 Results. Filtering of Drops Containing Cell Lysate from Burst Size Distributions.
(PDF)

S9 Results. Oseltamivir (OST) treatment of IAV Infected Cells During Drop Infections.
(PDF)

S10 Results. Modeling IAV Burst Size Distributions from Droplet RNA Concentration.
(PDF)

S1 Fig. Comparison of IAV M Gene Abundance During Bulk and Drop Infections. IAV strains (H1N1 and H3N2) were separately used to infect A549 cells under bulk cell culture and single cell encapsulation in 100 μm diameter microfluidic drops. M gene abundance (copies/μL) was measured using a bulk RT-qPCR assay from the supernatant of both bulk and drop infections at 0 and 18 hpi. Experimental details can be found in **S1 Results**. Each bar represents the pooled data from three replicate experiments. Error bars represent one standard deviation.
(TIF)

S2 Fig. Sigmoidal Curve Fitting using PCR Efficiency (SCF-E). A four-parametric sigmoid function was used to model decreasing PCR efficiency (E_N) at each cycle number (N) (Eq. S3). This model was fitted to discontinuous fluorescence measurements (F_N) of a known RNA template concentration ($C_{RNA,\ ref}$) to generate a continuous reference amplification curve (Eq. S4 and S5). Reference curves used in burst size experiments were generated with $C_{RNA,\ ref} = 1.71 \times 10^2$ RNA (cpd). In the SCF-E model, E_N is the PCR efficiency at cycle N; E_{max} is the maximum PCR efficiency at cycle number 1 and ranges from of 0.9 to 1; $E_{\min}$ is the minimum PCR efficiency at cycle number 40 and ranges from 0 to 0.1; $N_{0.5}$ is the cycle number at which E_N equals 0.5; and k is the shape parameter of the curve. As k increases, the shape of the curve flattens, and as k decreases, the shape of the curve becomes steeper. Thirty values of each of the four efficiency parameters ($E_{\max}$, $E_{\min}$, $N_{0.5}$, and k) were used to test $8.1{\times}10^5$ (30^4) curve fits to experimental fluorescence measurements. The set of efficiency parameters yielding the best fit were used to construct the SCF-E reference curve (solid blue curve).
(TIF)

S3 Fig. Translating a Single SCF-E Reference Curve into 1000 Virtual Amplification Curves. ACLs consist of virtual amplification curves that share the same $E_{\max}$, $E_{\min}$, and k parameters as the reference curve but have unique $N_{0.5}$ values. This is done by substituting $N_{0.5,\ ref}$ from Eq. S5 with $N_{0.5,\ virt}$ in Eq. S7. We input 1000 evenly spaced $N_{0.5,\ virt}$ values between cycle numbers $N = 1$ to 40, resulting in 1000 virtual amplification curves, each associated with a unique theoretical RNA concentration ($C_{RNA,\ virt}$). Virtual curves with $N_{0.5,\ virt} < N_{0.5,\ ref}$ (leftmost pink solid curve) correspond to higher RNA concentrations compared to the reference curve, whereas curves with $N_{0.5,\ virt} > N_{0.5,\ ref}$ (rightmost yellow solid curve) represent lower RNA concentrations than the reference (middle blue solid curve). The efficiency curve for reference and virtual curves is illustrated as a dashed blue line.
(TIF)

S4 Fig. Calculating RNA Template Concentration of the Virtual Amplification Curves. We relate the known RNA concentration of the reference curve ($C_{RNA,\ ref}$) to the unknown RNA concentrations of the virtual curves ($C_{RNA,\ virt}$) based on their position on the x-axis (cycle number, N). In Eq. S10, we define a relationship between the known cycle numbers where PCR efficiency is equal to 0.5 for the reference ($N_{0.5,\ ref}$) and virtual ($N_{0.5,\ virt}$) curves and their

respective RNA concentrations (C_{RNA}). By rearranging Eq. S10, we can calculate the unknown $C_{RNA,\ virt}$ using Eq. S12. To perform the calculation, the slope of the line (m) in Eq. S13 is determined using a PCR efficiency constant (E_N). As indicated by Eq. S13, an increase in E_N leads to a decrease in m (dotted black line), expanding the range of $C_{RNA,\ virt}$ in the ACL (from solid triangle to open triangle, and from solid square to open square). This relationship is visually illustrated here with Eq. S8.
(TIF)

S5 Fig. Constructing a Standard Curve from the ACL. (A) Construction of an amplification curve library (ACL) (S3 Fig). Three $C_{RNA,\ virt}$ within the ACL are displayed here (pink, blue, yellow), with the leftmost pink curve representing a higher $C_{RNA,\ virt}$ compared to the rightmost yellow curve. **(B)** ACL standard curves (solid line, dashed line, dotted line) are constructed from the F_N of $C_{RNA,\ virt}$ (pink, blue, yellow) at a particular cycle number (15, 20, 25). The cycle number chosen for the ACL standard curve corresponds to the cycle number at which unknown RNA concentrations are sampled. In this method, only virtual curves in the exponential or linear phase at the selected cycle number can be included in the standard curve (grey shaded regions in **A** and **B**). This is because the early amplification and plateau regions of multiple virtual curves may overlap at a single cycle number. To ensure that only useable regions are selected for the standard curve, we establish threshold values based on the fluorescence of the reference curve. The upper threshold is set at 1 standard deviation below the F_N at cycle 40, and the lower threshold is set at the 99th percentile of the F_N at cycle 1. For example, as amplification proceeds for a particular $C_{RNA,\ virt}$ (pink curve) in **A**, the F_N increases and intersects cycles 15, 20, and 25 at the three points (squares). However, only two of those points are within the valid range of the ACL standard curves in **B.**
(TIF)

S6 Fig. Comparison of SCF-E and SCF Methods of Fitting PCR Amplification Curves. Fluorescence intensities (ΔR_N) of bulk RT-qPCR amplification of six concentrations of IAV M gene RNA (S2 Table, 10-fold dilutions from 2.62×10^4 to 2.62×10^9 copies/μL) were measured at $N = 1$ to 40. The concentration of each dilution decreases from left to right when viewing the amplification curves. **(A)** Amplification curves generated with the SCF-E model (dotted gray curves) had an average fit of $R^2 = 0.9996$ across all RNA concentrations (S3 Table). **(B)** Amplification curves generated with the SCF model (dotted gray curves) had an average fit of $R^2 = 0.9975$ across all RNA concentrations (S3 Table). **(C)** For a single M gene RNA concentration (2.62×10^7 copies/μL), the SCF-E model provided a better fit ($R^2 = 0.9996$) than SCF ($R^2 = 0.9972$). The difference in fit is most noticeable at the lower ($\Delta R_N < 1$) and higher ($\Delta R_N > 4$) fluorescence values, representing the exponential and plateau phases of PCR amplification (magnified insets). All error bars represent one standard deviation from the average fluorescence measured across three technical replicates.
(TIF)

S7 Fig. Setting a PCR Efficiency Constant (EN) for the ACL. To calculate E_N at $N = N_{0.5}–3$, we determined $N_{0.5}$ of the SCF-E reference curve, which corresponds to N where $E_N = 0.5$ **(A to B)**. $N_{0.5}–3$ represents three cycles prior to $N_{0.5}$ **(C)**. We recorded the E_N at $N_{0.5}–3$ **(D)**. When $N = N_{0.5}–3$ is inserted into Eq. S3, the term $N_{0.5}$ cancels out, resulting in a constant value of -3 for ($N–N_{0.5}$). The E_N calculated in this step is used in Eq. S13 to determine m.
(TIF)

S8 Fig. Linear regression of measured to expected RNA concentrations from S4 Table. We compared cycle numbers (N) at which PCR efficiency (E_N, Eq. S3) of the reference curve is used to calculate the constant m (Eq. S13), for assigning RNA template concentrations to

virtual curves in the ACL. We tested the cycle number at which PCR efficiency is equal to 0.5 ($N_{0.5}$) and the five cycles preceding it ($N = N_{0.5}–1$ *to* $N = N_{0.5}–5$), to construct six ACLs for comparison. From each ACL, a standard curve was constructed at PCR cycle number 20. We evaluated the dynamic range of each standard curve in S4 Table. The ACL standard curve with the widest dynamic range and the highest degree of accuracy (R^2) between measured ($C_{RNA, ACL}$) and expected ($C_{RNA, expected}$) M gene RNA concentrations was built using E_N at $N = N_{0.5}–3$. For this ACL, the dynamic range of the standard curve was 4.70 $\log_{10}$ RNA copies/μL (S4 Table, dynamic range of 2.62×10^4 to 1.31×10^9 copies/μL) with $R^2 = 0.990$ between $C_{RNA, ACL,}$ and $C_{RNA, expected}$. Therefore, the E_N at $N = N_{0.5}–3$ was used to construct all amplification curve libraries in this work.
(TIF)

S9 Fig. Linear regression of measured to expected RNA concentrations from S5 Table. We compared the dynamic range of ACL standard curves from different cycle numbers ($N = 17$ to 23) to the Ct standard curve. Standard curves were used to convert fluorescence to template concentration for a known dilution series of IAV M gene RNA (S2 Table, three replicates of 21 concentrations between 2.62×10^4 to 2.62×10^9 copies/uL). Standard curve dynamic range and degree of accuracy (R^2) between measured ($C_{RNA, ACL}$) and expected ($C_{RNA, expected}$) M gene RNA concentrations were evaluated in S5 Table.
(TIF)

S10 Fig. Variability in M gene RNA Concentrations Measured with ACL Standard Curves by PCR Cycle Number. We examined variability between standard curves obtained from different ACL cycle numbers using two measures: **(A)** coefficient of variation (*CV* %) and **(B)** percent difference (%). Standard curves were constructed from ACL cycle numbers ($N = 17$ to 23, individual plots) and used to convert fluorescence to RNA concentration ($C_{RNA, ACL}$) across 23 IAV M gene RNA concentrations (S2 Table). **(A)** *CV* (%) is the ratio of the standard deviation to the mean $C_{RNA, ACL}$, and it quantifies the variability *within* replicates of the same M gene RNA concentration. A linear regression (dotted line) was conducted to compare *CV* (%) across M gene RNA concentrations. The results revealed a weak linear relationship ($R^2 <$ 0.28) for all cycle numbers. **(B)** Percent difference (%) measures the variability *between* $C_{RNA, ACL}$ and the expected RNA concentration ($C_{RNA, expected}$). Similarly, a weak relationship ($R^2 <$ 0.36) was observed between the starting RNA concentration and percent difference (%) for all cycle numbers. To assess the variability in ACL standard curves, a comparison was made with *Ct* standard curves (A and B, first upper left plot), for *CV* (%) and percent difference (%). The analysis indicated no linear response for *CV* (%) ($R^2 = 0.0236$) and a weak linear response for percent difference (%) ($R^2 = 0.2477$) across the M gene RNA concentrations.
(TIF)

S11 Fig. Validating ACL Standard Curves for Multiple Template Sequences. Standard curves constructed using the **(A)** ACL method or **(B)** Ct method were used to convert fluorescence to template concentration ($C_{RNA/DNA, ACL}$) for a known dilution series of three different template sequences: IAV M gene RNA (green/gray dots, ranging from 2.62×10^4 to 2.92×10^9 copies/μL), human MYCN gene DNA (blue squares, ranging from 1.88×10^0 to 1.88×10^3 copies/μL), and soybean Lectin endogene DNA (red circles, ranging from 6.4×10^0 to 4.0×10^3 copies/μL). Each dilution of the templates had varying numbers of technical replicates: 3 for the M gene, 94 for the MYCN, and 18 for the Lectin dataset. The M gene ACL was built with a 10^7 copies/μL reference curve fitted with either the SCF-E (green dots) or SCF (grey dots) model. The standard curve from the M gene ACL was constructed at a single cycle number ($N = 20$). The MYCN libraries were built with a 1,875 copies/μL reference curve, and

the Soybean ACL was built with an 800 copies/μL reference curve, both fitted using the SCF-E model. The standard curves from these ACLs were constructed at $N = 27$ (MYCN) and $N = 29$ (Lectin). An R^2 value quantified the linear response between measured $C_{RNA/DNA,\ ACL}$ to expected concentrations ($C_{RNA/DNA,\ expected}$) across template concentrations. Dotted lines indicate a 2-fold change from $C_{RNA/DNA,\ expected}$, and error bars represent one standard deviation from the mean.
(TIF)

S12 Fig. Droplet Quantitative PCR (dqPCR) of the IAV M gene. (A) SCF-E reference amplification curve generated for 1.71×10^2 cpd of IAV M gene RNA amplified in 50 μm drops. Drop fluorescence intensity (ΔR_N) was detected at PCR cycle numbers $N = 1$, 13, 16, 19, 22, 25, 28, 40 and displayed as a mean (black dots) and one standard deviation (error bars) (≈1000 or more drops per cycle number). The estimated PCR efficiency (E_N) curve (dashed blue line) used during SCF-E produced a well-fitting reference amplification curve (dotted red line, $R^2 = 0.9995$). **(B)** dqPCR was validated for a single concentration of M gene RNA (1.71×10^2 cpd) across three biological replicates. The samples were collected at various cycle numbers: rep 1 (red squares) at $N = 19$ to 21 and rep 2 (blue circles) and rep 3 (green triangles) at $N = 19$ to 22. Drop fluorescence was converted to M gene RNA concentration ($C_{RNA,\ ACL}$) within a 2-fold change of the expected mean (dotted lines) for all cycle numbers. Dashed lines indicate a fixed effects model with upper and lower bounds set as the 95% confidence intervals. **(C)** Representative epifluorescence images of 50 μm diameter drops at PCR cycle numbers 1, 19, and 40.
(TIF)

S13 Fig. SCF-E Reference Amplification Curves used in Fig 3B. Three known concentrations of M gene RNA, 1.71×10^1 cpd (red dots), 1.71×10^2 cpd (blue dots), and 1.71×10^3 cpd (green dots), were amplified in 50 μm drops. Drop fluorescence (ΔR_N) was detected at multiple PCR cycle numbers (S6 Table, note that cycles 19–21 in the 10 cpd group were below background and not included in the table) and converted to M gene cpd using dqPCR (Fig 3B). For each experiment run on different days, a new SCF-E reference curve was generated from the amplification of 1.71×10^2 cpd (red dotted line, blue dashed line, green solid line).
(TIF)

S14 Fig. Using dqPCR to Quantify Three IAV M gene RNA Concentrations at Multiple Cycle Numbers. Three known IAV M gene RNA concentrations, 1.17×10^1 cpd (10 cpd, red circles), 1.17×10^2 cpd (100 cpd, blue crosses), and 1.17×10^3 cpd (1000 cpd, green squares) were amplified in 50 μm drops. Drop fluorescence (ΔR_N) was detected at multiple PCR cycle numbers (S6 Table) and converted to M gene cpd ($C_{RNA,\ ACL}$) using dqPCR. $C_{RNA,\ ACL}$ measurements were pooled together and displayed as a single distribution in Fig 3B. Dashed lines indicate the expected M gene cpd, and dotted lines represent a 2-fold change from the expected mean.
(TIF)

S15 Fig. Analysis of Variance (ANOVA) for dqPCR of Three IAV M Gene RNA Concentrations. Three known IAV M gene RNA concentrations, 1.17×10^1 cpd (10 cpd plot), 1.17×10^2 cpd (100 cpd plot), and 1.17×10^3 cpd (1000 cpd plot), were amplified in 50 μm drops. Drop fluorescence (ΔR_N) was detected at multiple PCR cycle numbers (S6 Table) and converted to M gene cpd ($C_{RNA,\ ACL}$) using the dqPCR model. $C_{RNA,\ ACL}$ measurements were pooled together and displayed as a single distribution in Fig 3B. To examine the contribution of error within individual PCR cycle numbers to overall error in the pooled distributions, we performed an Analysis of Variance (ANOVA) test on a random sample of 100 drops from each group. **(A)** The 10 cpd group, quantified at PCR cycle numbers 22–25, had mean $C_{RNA,}$

$_{ACL}$ values that differed by cycle number (ANOVA p-value <0.05). This variability due to cycle number accounted for 8% of the total error in the pooled distribution. **(B)** The 100 cpd group, quantified at PCR cycle numbers 19–22, had mean $C_{RNA,\ ACL}$ values that differed by cycle number (ANOVA p-value <0.05). This variability due to cycle number accounted for 59% of the total error in the pooled distribution. **(C)** The 1000 cpd group, quantified at PCR cycle numbers 16–22, had mean $C_{RNA,\ ACL}$ values that differed by cycle number (ANOVA p-value <0.05). This variability due to cycle number accounted for 7% of the total error in the pooled distribution.
(TIF)

S16 Fig. Linear Mixed Effects Model (LME) for dqPCR of Three IAV M Gene RNA Concentrations. Three known IAV M gene RNA concentrations, 1.17×10^1 cpd (10 cpd, red circle), 1.17×10^2 cpd (100 cpd, blue cross), and 1.17×10^3 cpd (1000 cpd, green triangle), were amplified in 50 μm drops. Drop fluorescence (ΔR_N) was detected at multiple PCR cycle numbers (S6 Table) and converted to M gene cpd ($C_{RNA,\ ACL}$). $C_{RNA,\ ACL}$ measurements were pooled together and displayed as a single distribution in Fig 3B. To further examine the contribution of error within individual PCR cycle numbers to overall error in the pooled distributions, we calculated the percent difference (%) (Eq. S14) between the measured $C_{RNA,\ ACL}$ values and their expected M gene cpd ($C_{RNA,\ expected}$), across cycle numbers. This analysis informed the validity of pooling conversions measured from multiple PCR cycles in our burst size experiments. Percent differences were compared using a linear mixed effects model (LME) with a 95% confidence interval (CI) (dotted lines). In the LME, we set cycle as the fixed effect variable and $C_{RNA,\ expect}$ as the random effect variable. Both variables were log-transformed during the model fit, as replication cycles correspond to RNA copy numbers on a log scale. The fixed effects coefficients described a negative relationship between the slope and percent difference from the expected RNA concentration of -0.85% per cycle number (p-value <0.05, SE = 0.01). While the p-value indicates significance, the magnitude of the relationship is small (~1%/cycle), maintaining that it is valid to pool dqPCR data from different cycle numbers.
(TIF)

S17 Fig. Using dqPCR to Quantify IAV M gene RNA Concentration, Pooled Cycle Numbers. Three known IAV M gene RNA concentrations, 1.17×10^1 cpd, 1.17×10^2 cpd, 1.17×10^3 cpd, were amplified in 50 μm drops. Drop fluorescence (ΔR_N) was detected at multiple PCR cycle numbers (S6 Table) and converted to M gene cpd ($C_{RNA,\ ACL}$) using dqPCR. $C_{RNA,\ ACL}$ measurements were pooled together and plotted against the expected M gene cpd ($C_{RNA,\ expected}$). Two biological replicates (black circle and green triangle) are shown. The solid line represents a 1:1 relationship between $C_{RNA,\ ACL}$ and $C_{RNA,\ expected}$, dashed lines indicate a 2-fold change from the expected mean.
(TIF)

S18 Fig. β-actin mRNA abundance increases with cell concentration. A bulk RT-qPCR assay was performed to amplify β-actin mRNA from lysed A549 cells after 24 hr of incubation. There was a strong linear relationship between β-actin and cell concentration (dashed black line, $R^2 = 0.995$), validating its use as a reliable marker for the detection of intracellular lysate in drops during our burst size experiments.
(TIF)

S19 Fig. Multiplexed dqPCR for Detection of IAV M gene and Cellular β-actin. (A) Both M gene RNA (red) and β-actin mRNA (black) were detected by multiplexed dqPCR from H1N1 infection in drops (n = 2000). Data is presented in ascending order of RNA concentration,

prior to β-actin filtering (indicated by the dashed black line). **(B)** Showing a zoomed-in region of (A), with drop indexes 1900 to 2000.
(TIF)

S20 Fig. SCF-E Reference Curves for Burst Size Replicate Experiments. For each burst size replicate experiment (S8 Table), a new reference amplification curve was constructed for dqPCR (S9 Table). To generate these curves, a known concentration of M gene RNA (1.71×10^2 cpd) was amplified in 100 μm drops. Drop fluorescence intensity (ΔR_N) was measured at PCR cycle numbers $N = 1$, 16, 19, 22. 25, 28, 31 and 40. The ΔR_N of drops from each cycle number are shown as the mean (black dots) with error bars representing one standard deviation. SCF-E was used to create a continuous reference amplification curve (orange dashed line) from discontinuous ΔR_N measurements, using an estimate for the PCR efficiency (blue dashed line).
(TIF)

S21 Fig. Oseltamivir (OST) treatment of IAV infected cells during drop infections. MDCK cells infected with IAV H1N1 were encapsulated into 100 μm diameter microfluidic drops. During encapsulation, infected cells were either suspended in standard droplet infection media (+H1N1 -OST) or in droplet infection media treated with 10 μM concentration of oseltamivir (+H1N1 +OST). Mock infected cells encapsulated in standard infection media (-H1N1 –OST) were used as a negative control. M gene abundance (copies / μL) was measured using a bulk RT-qPCR assay from the supernatant of broken drops sampled at 0 and 18 hpi. Each bar represents the pooled data from three technical replicates. Significance stars (*** $p < 0.001$ and **** $p < 0.0001$) obtained by a two-sample Student's t-test.
(TIF)

S22 Fig. Best Fit Model to Estimate IAV Burst Size Distributions from S10 Table. We used a simulation-based approach to estimate the IAV burst distributions shown in Fig 4D. We considered three distribution models: lognormal (solid gray curve), Poisson (dashed gray curve), and negative-binomial distribution (solid blue or red curve, corresponding to H3N2 and H1N1 distributions, respectively). The parameters for each distribution are provided in S10 Table. Based on the analysis, burst sizes were found to best fit a negative binomial distribution.
(TIF)

S23 Fig. Cell Loading During Burst Size Replicate Experiments. (A) Images of 100 μm drops re-injected into a microfluidic device were captured at the beginning and end of each burst size replicate experiment to monitor cell loading. Drops containing a single cell are indicated by blue arrows. **(B)** Across all H3N2 and H1N1 replicate experiments, cell loading was determined by counting the percentage of drops containing cells (mean = 25% ± 4%). Cell loading in drops was assumed to follow a Poisson distribution. For a population of drops with 25% containing cells, the estimated Poisson mean (λ) is 0.29 cells/drop. In this case, ≈75% of drops are empty, ≈21.5% of drops contain one cell, ≈3% contain two cells, and ≈0.5% of drops contain three or more cells. The bar graph displays the drop counts for the replicates from left to right, which are 753, 679, 538, 719, 774, and 582.
(TIF)

S1 Table. RT-qPCR Template and Primer Sequences. All sequences are listed in the 5' to 3' direction. The M gene template control was constructed as linear dsDNA containing a T7 promoter (underlined), forward and reverse primer binding sites (*italicized*), a probe binding site (**bold**), and a partial M gene sequence. The M gene dsDNA was *in vitro* transcribed prior to

use as a positive control. The M gene TaqMan probe carried a FAM fluorophore with a BHQ1 quencher. The β-actin template control was constructed as a dsDNA plasmid containing forward and reverse primer binding sites (*italicized*), and a probe binding site (**bold**), and a partial β-actin mRNA sequence. The β-actin TaqMan probe carried a Cy5 fluorophore with a BHQ2 quencher.
(TIF)

S2 Table. IAV M Gene RNA Concentrations Amplified by Bulk RT-qPCR and Used to Validate dqPCR.
(TIF)

S3 Table. SCF-E and SCF Curve Fitting Parameters for S6 Fig.
(TIF)

S4 Table. Dynamic Range of ACL Standard Curves Built with Different EN Constants. We compared cycle numbers (N) at which PCR efficiency (E_N) of the reference curve is used to calculate the constant m (and E_N). This constant determines the RNA template concentration assigned to virtual curves in the ACL. E_N is calculated using Eq. S3 and m is calculated using Eq. S13. We used the cycle number at which PCR efficiency is equal to 0.5 ($N_{0.5}$) and the five cycles preceding it ($N = N_{0.5}-1$ *to* $N = N_{0.5}-5$), to construct six ACLs for comparison. From each ACL, a standard curve was generated at PCR cycle number 20. This standard curve was then used to convert fluorescence (F_N) to template concentration (C_{RNA}) for a known dilution series of M gene RNA (S2 Table, three replicates of 21 concentrations between 2.62×10^4 to 2.62×10^9 copies/μL). The minimum ($C_{RNA,\ min}$) and maximum ($C_{RNA,\ max}$) concentrations in the dilution series, whose measured RNA concentration fell within a 2-fold change of the expected concentrations, marked the dynamic range of each standard curve. The R^2 value quantifies the linear regression (S8 Fig) between the measured ($C_{RNA,\ ACL}$) and expected ($C_{RNA,\ expected}$) RNA concentrations. The ACL built with E_N at $N = N_{0.5}-3$ produced a cycle 20 standard curve with the highest dynamic range and the greatest degree of accuracy (R^2). Therefore, an offset of -3 cycles was used to construct all ACLs in this work.
(TIF)

S5 Table. Dynamic Range of ACL Standard Curves from Different ACL Cycle Numbers. We compared the dynamic range of standard curves created using the amplification curve library (ACL) method to that of the cycle threshold (*Ct*) method. Standard curves were used to convert fluorescence to template concentration for a known dilution series of IAV M gene RNA (S2 Table, three replicates of 21 concentrations between 2.62×10^4 to 2.62×10^9 copies/μL). In the ACL method, multiple standard curves were constructed from individual cycle numbers ($N = 17$ to 23). On the other hand, in the *Ct* method, a single standard curve is built from multiple *Ct* values, with one *Ct* value corresponding to each concentration of M gene RNA. The dynamic range of each standard curve was determined by identifying the minimum ($C_{RNA,\ min}$) and maximum ($C_{RNA,\ max}$) concentrations within the dilution series that yielded measured RNA concentration falling within a 2-fold change of the expected concentrations. The R^2 value quantifies the linear response (S9 Fig) between the measured ($C_{RNA,\ ACL}$) and expected ($C_{RNA,\ expected}$) RNA concentrations.
(TIF)

S6 Table. Using dqPCR to Quantify IAV M Gene RNA Concentrations at Multiple Cycle Numbers. Three known IAV M gene RNA concentrations, 1.71×10^1 cpd (10^1), 1.71×10^2 cpd (10^2), and 1.71×10^3 cpd (10^3), were amplified in 50 μm drops. Drop fluorescence (ΔR_N) was detected at multiple PCR cycle numbers and converted to M gene cpd ($C_{RNA,\ ACL}$) using

dqPCR. $C_{RNA,\ ACL}$ measurements from different cycle numbers were pooled together and presented as a single distribution in Fig 3B. To assess the variability between individual cycle numbers and the pooled distribution, we analyzed the standard deviation of the mean $C_{RNA,\ ACL}$ and the coefficient of variation (CV %) across the number of sampled drops (n) at each N.
(TIF)

S7 Table. Maximum Likelihood Estimates for a Gaussian Mixture Model (GMM) Used to Extract Individual Distributions from Mixed Droplet Detection Data. To use dqPCR for measuring IAV burst size in a heterogeneous population of single-cell infections, we validated the method's ability to isolate individual distributions from a sample containing multiple M gene RNA concentrations. We applied a Gaussian mixture model (GMM) on dqPCR data obtained from three known IAV M gene RNA concentrations: 1.17×10^1 cpd, 1.17×10^2 cpd, and 1.17×10^3 cpd, mixed together in a single sample (data shown in Fig 3E). Model assumptions are described in **S7 Results**. In the model, x_i represents the number of M gene RNA copies (cpd). Each Gaussian has a different weight (w_i) corresponding to the probability that a random drop in the mixture has x_i copies, a predicted mean (B_i, representing measurement bias) and variance (σ_i^2, representing measurement noise).
(TIF)

S8 Table. Burst Size Replicate Experiments. IAV burst sizes were measured from thousands of individual drops sampled at cycle numbers N = 16, 19, 22, 25, and 28 in three biological replicates. The measurements were used to determine the number of IAV M gene RNA and cellular β-actin mRNA (cpd). Drops containing high β-actin cpd (S19 Fig) were assumed to contain non-packaged, intracellular, M gene RNA and were subsequently filtered from the final burst size distributions (Filtered Drops, S8 Results). We present the average measured burst size for each replicate, along with the interquartile range (IQR) that represents the lower 25^{th} and upper 75^{th} percentile of the distributions. The heterogeneity of each distribution was examined using the Gini coefficient (G). If $G = 0$, the distribution is completely homogeneous, meaning that each cell produces the same number of viruses. Conversely, if $G = 1$, the distribution is completely heterogeneous, where one cell produces all the viruses. Thus, a higher G corresponds to a more heterogeneous distribution.
(TIF)

S9 Table. dqPCR Parameters Used in Burst Size Replicate Experiments. The parameters used to create reference standard curves using SCF-E (S20 Fig, orange dotted curves) for each burst size experiment.
(TIF)

S10 Table. Best Fit Model to Estimate IAV Burst Size in Fig 4D. We used a simulation-based approach to estimate the IAV burst distributions shown in Fig 4D. We considered three possible distributions: lognormal, Poisson, and negative-binomial, all with unknown parameters. First, we simulated the viral burst size based on the assumed distribution. For each simulated value of burst size x, we introduced measurement noise by assuming a log-normal distribution. The mean of the log-normal distribution was determined by the bias function $B(x)$, and the standard deviation was determined by $\sigma(x)$. Next, we computed a density function using a kernel density estimation with a Gaussian kernel to represent the resulting distribution. We then used this density function to calculate the log-likelihood of the observed data. We estimated the parameters for each distribution and for each dataset (H3N2 or H1N1) by maximizing the likelihood of observations reported in Fig 4D.
(TIF)

Acknowledgments

We thank Humberto Sanchez for help with protocols and Carter Hoffman for helpful discussions.

Author Contributions

Conceptualization: Geoffrey K. Zath, Ruian Ke, Christopher B. Brooke, Connie B. Chang.

Data curation: Geoffrey K. Zath, Mallory M. Thomas.

Formal analysis: Geoffrey K. Zath, Mallory M. Thomas, Steven Sanche, Ruian Ke.

Funding acquisition: Ruian Ke, Christopher B. Brooke, Connie B. Chang.

Investigation: Geoffrey K. Zath, Mallory M. Thomas, Emma K. Loveday, Dimitri A. Bikos.

Methodology: Geoffrey K. Zath, Mallory M. Thomas, Emma K. Loveday, Dimitri A. Bikos, Connie B. Chang.

Project administration: Connie B. Chang.

Resources: Ruian Ke, Christopher B. Brooke, Connie B. Chang.

Software: Geoffrey K. Zath.

Supervision: Connie B. Chang.

Validation: Geoffrey K. Zath, Mallory M. Thomas.

Visualization: Geoffrey K. Zath, Mallory M. Thomas.

Writing – original draft: Geoffrey K. Zath.

Writing – review & editing: Geoffrey K. Zath, Mallory M. Thomas, Connie B. Chang.

References

1. Burnet FM. A Method for the Study of Bacteriophage Multiplication in Broth. Br J Exp Pathol. 1929 Apr; 10(2):109.
2. Delbrück M. The Burst Size Distribution in the Growth of Bacterial Viruses (Bacteriophages). J Bacteriol. 1945 Aug; 50(2):131–5. PMID: 16560982
3. DeLong JP, Al-Sammak MA, Al-Ameeli ZT, Dunigan DD, Edwards KF, Fuhrmann JJ, et al. Towards an integrative view of virus phenotypes. Nat Rev Microbiol. 2021 Sep; https://doi.org/10.1038/s41579-021-00612-w PMID: 34522049
4. Jones JE, Le Sage V, Lakdawala SS. Viral and host heterogeneity and their effects on the viral life cycle. Nat Rev Microbiol. 2021 Apr; 19(4):272–82. https://doi.org/10.1038/s41579-020-00449-9 PMID: 33024309
5. 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. Nat Commun. 2015 Nov; 6:8938. https://doi.org/10.1038/ncomms9938 PMID: 26586423
6. Kupke SY, Ly LH, Börno ST, Ruff A, Timmermann B, Vingron M, et al. Single-Cell Analysis Uncovers a Vast Diversity in Intracellular Viral Defective Interfering RNA Content Affecting the Large Cell-to-Cell Heterogeneity in Influenza A Virus Replication. Viruses. 2020 Jan; 12(1). https://doi.org/10.3390/v12010071 PMID: 31936115
7. Russell AB, Trapnell C, Bloom JD. Extreme heterogeneity of influenza virus infection in single cells. Elife. 2018 Feb; 7. https://doi.org/10.7554/eLife.32303 PMID: 29451492
8. Sun J, Vera JC, Drnevich J, Lin YT, Ke R, Brooke CB. Single cell heterogeneity in influenza A virus gene expression shapes the innate antiviral response to infection. PLoS Pathog. 2020 Jul; 16(7): e1008671. https://doi.org/10.1371/journal.ppat.1008671 PMID: 32614923
9. Bacsik DJ, Dadonaite B, Butler A, Greaney AJ, Heaton NS, Bloom JD. Influenza virus transcription and progeny production are poorly correlated in single cells. bioRxiv. 2022.

10. Jacobs NT, Onuoha NO, Antia A, Steel J, Antia R, Lowen AC. Incomplete influenza A virus genomes occur frequently but are readily complemented during localized viral spread. Nat Commun. 2019 Aug; 10(1):3526. https://doi.org/10.1038/s41467-019-11428-x PMID: 31387995
11. Brooke CB, Ince WL, Wrammert J, Ahmed R, Wilson PC, Bennink JR, et al. Most influenza a virions fail to express at least one essential viral protein. J Virol. 2013 Mar; 87(6):3155–62. https://doi.org/10.1128/JVI.02284-12 PMID: 23283949
12. Brooke CB. Population Diversity and Collective Interactions during Influenza Virus Infection. J Virol. 2017 Nov; 91(22). https://doi.org/10.1128/JVI.01164-17 PMID: 28855247
13. Schulte MB, Andino R. Single-cell analysis uncovers extensive biological noise in poliovirus replication. J Virol. 2014 Jun; 88(11):6205–12. https://doi.org/10.1128/JVI.03539-13 PMID: 24648454
14. Vignuzzi M, López CB. Defective viral genomes are key drivers of the virus–host interaction. Nature Microbiology. 2019 Jun; 4(7):1075–87. https://doi.org/10.1038/s41564-019-0465-y PMID: 31160826
15. Guo MT, Rotem A, Heyman JA, Weitz DA. Droplet microfluidics for high-throughput biological assays. Lab Chip. 2012 Jun; 12(12):2146–55. https://doi.org/10.1039/c2lc21147e PMID: 22318506
16. Fischer AE, Wu SK, Proescher JBG, Rotem A, Chang CB, Zhang H, et al. A high-throughput drop microfluidic system for virus culture and analysis. J Virol Methods. 2015 Mar; 213:111–7. https://doi.org/10.1016/j.jviromet.2014.12.003 PMID: 25522923
17. Tao Y, Rotem A, Zhang H, Cockrell SK, Koehler SA, Chang CB, et al. Artifact-Free Quantification and Sequencing of Rare Recombinant Viruses by Using Drop-Based Microfluidics. Chembiochem. 2015 Oct; 16(15):2167–71. https://doi.org/10.1002/cbic.201500384 PMID: 26247541
18. Zhang H, Cockrell SK, Kolawole AO, Rotem A, Serohijos AWR, Chang CB, et al. Isolation and Analysis of Rare Norovirus Recombinants from Coinfected Mice Using Drop-Based Microfluidics. J Virol. 2015 Aug; 89(15):7722–34. https://doi.org/10.1128/JVI.01137-15 PMID: 25972549
19. Rotem A, Serohijos AWR, Chang CB, Wolfe JT, Fischer AE, Mehoke TS, et al. Evolution on the Biophysical Fitness Landscape of an RNA Virus. Mol Biol Evol. 2018 Oct; 35(10):2390–400. https://doi.org/10.1093/molbev/msy131 PMID: 29955873
20. Tao Y, Rotem A, Zhang H, Chang CB, Basu A, Kolawole AO, et al. Rapid, targeted and culture-free viral infectivity assay in drop-based microfluidics. Lab Chip. 2015 Sep; 15(19):3934–40. https://doi.org/10.1039/c5lc00556f PMID: 26304791
21. Loveday EK, Sanchez HS, Thomas MM, Chang CB. Single-Cell Infection of Influenza A Virus Using Drop-Based Microfluidics. Microbiol Spectr. 2022 Sep;e0099322. https://doi.org/10.1128/spectrum.00993-22 PMID: 36125315
22. Hindson BJ, Ness KD, Masquelier DA, Belgrader P, Heredia NJ, Makarewicz AJ, et al. High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. Anal Chem. 2011 Nov; 83(22):8604–10. https://doi.org/10.1021/ac202028g PMID: 22035192
23. Hatch AC, Fisher JS, Tovar AR, Hsieh AT, Lin R, Pentoney SL, et al. 1-Million droplet array with wide-field fluorescence imaging for digital PCR. Lab Chip. 2011 Nov; 11(22):3838–45. https://doi.org/10.1039/c1lc20561g PMID: 21959960
24. Beer NR, Hindson BJ, Wheeler EK, Hall SB, Rose KA, Kennedy IM, et al. On-chip, real-time, single-copy polymerase chain reaction in picoliter droplets. Anal Chem. 2007 Nov; 79(22):8471–5. https://doi.org/10.1021/ac701809w PMID: 17929880
25. Beer NR, Wheeler EK, Lee-Houghton L, Watkins N, Nasarabadi S, Hebert N, et al. On-chip single-copy real-time reverse-transcription PCR in isolated picoliter droplets. Anal Chem. 2008 Mar; 80(6):1854–8. https://doi.org/10.1021/ac800048k PMID: 18278951
26. Kiss MM, Ortoleva-Donnelly L, Beer NR, Warner J, Bailey CG, Colston BW, et al. High-throughput quantitative polymerase chain reaction in picoliter droplets. Anal Chem. 2008 Dec; 80(23):8975–81. https://doi.org/10.1021/ac801276c PMID: 19551929
27. Prakash R, Pabbaraju K, Wong S, Wong A, Tellier R, Kaler KVIS. Multiplex, Quantitative, Reverse Transcription PCR Detection of Influenza Viruses Using Droplet Microfluidic Technology. Micromachines. 2014 Dec; 6(1):63–79.
28. Prakash R, Pabbaraju K, Wong S, Wong A, Tellier R, Kaler KVIS. Droplet Microfluidic Chip Based Nucleic Acid Amplification and Real-Time Detection of Influenza Viruses. J Electrochem Soc. 2014 Jan; 161(2):B3083–93.
29. Hajji I, Serra M, Geremie L, Ferrante I, Renault R, Viovy JL, et al. Droplet microfluidic platform for fast and continuous-flow RT-qPCR analysis devoted to cancer diagnosis application. Sens Actuators B Chem. 2020 Jan; 303:127171.
30. Hayes CJ, Dalton TM. Microfluidic droplet-based PCR instrumentation for high-throughput gene expression profiling and biomarker discovery. Biomol Detect Quantif. 2015 Jun; 4:22–32. https://doi.org/10.1016/j.bdq.2015.04.003 PMID: 27077035

31. Loveday EK, Zath GK, Bikos DA, Jay ZJ, Chang CB. Screening of Additive Formulations Enables Off-Chip Drop Reverse Transcription Quantitative Polymerase Chain Reaction of Single Influenza A Virus Genomes. Anal Chem. 2021 Mar; 93(10):4365–73. https://doi.org/10.1021/acs.analchem.0c03455 PMID: 33635052

32. Martin BE, Harris JD, Sun J, Koelle K, others. Cellular co-infection can modulate the efficiency of influenza A virus production and shape the interferon response. PLoS. 2020; https://doi.org/10.1371/journal.ppat.1008974 PMID: 33064776

33. Mazutis L, Gilbert J, Ung WL, Weitz DA, Griffiths AD, Heyman JA. Single-cell analysis and sorting using droplet-based microfluidics. Nat Protoc. 2013 May; 8(5):870–91. https://doi.org/10.1038/nprot.2013.046 PMID: 23558786

34. Gini C. Measurement of Inequality of Incomes. Econ J. 1921 Mar; 31(121):124–5.

35. Carlo DD, Irimia D, Tompkins RG, Toner M. Continuous inertial focusing, ordering, and separation of particles in microchannels. Proceedings of the National Academy of Sciences. 2007; 104(48):18892–7. https://doi.org/10.1073/pnas.0704958104 PMID: 18025477

36. Abate AR, Chen CH, Agresti JJ, Weitz DA. Beating Poisson encapsulation statistics using close-packed ordering. Lab Chip. 2009 Sep; 9(18):2628–31. https://doi.org/10.1039/b909386a PMID: 19704976

37. Chen KY, Karuppusamy J, O'Neill MB, Opuu V, Bahin M, Foulon S, et al. High-throughput droplet-based analysis of influenza A virus genetic reassortment by single-virus RNA sequencing. Proc Natl Acad Sci U S A. 2023 Feb; 120(6):e2211098120. https://doi.org/10.1073/pnas.2211098120 PMID: 36730204

CHAPTER FIVE

DROPLET MICROFLUIDIC PLATFORM FOR SINGLE-CELL SEQUENCING OF MEASLES VIRUS DURING ONCOLYTIC VIROTHERAPY

Contribution of Authors and Co-Authors

Manuscript in Chapter 5

Author: Mallory Blackwood

Contributions: Designed studies, Performed Experiments, Analyzed data, Wrote Manuscript.

Co-Author: Ianko Iankov

Contributions: Designed studies, Performed Experiments.

Co-Author: Lydia Jordache

Contributions: Performed Experiments.

Co-Author: Casper Chung

Contributions: Designed studies, Performed Experiments.

Co-Author: Damon Trump

Contributions: Performed Experiments.

Co-Author: Eva Galanis

Contributions: PI on the Study, Acquired funding.

Co-Author: Connie Chang

Contributions: PI on the Study, Acquired funding, Designed studies, Wrote Manuscript.

Manuscript Information

Mallory Blackwood, Ianko Iankov, Lydia Jordache, Casper Chung, Damon Trump, Jonathan Louis, Connie Chang.

Biomicrofluidics

Status of Manuscript:
☒ Prepared for submission to a peer-reviewed journal
☐ Officially submitted to a peer-reviewed journal
☐ Accepted by a peer-reviewed journal
☐ Published in a peer-reviewed journal

AIP Publishing

Abstract

The RNA-stranded measles virus (MV) has a genome subject to frequent mutation, which provides selective advantages in diverse host cell environments. During *in vitro* passaging of the MV vaccine strain, random mutation in the P gene led to the development of a new attenuation phenotype with enhanced potential for cancer oncolytic therapy. Attenuated MVs specifically target cancer cells that overexpress the CD46 membrane receptor, infecting and killing these cells through syncytia formation, while simultaneously triggering a localized immune response at the tumor site. This represents a promising approach to cancer oncolytic therapy that is currently being evaluated in Phase I and II clinical trials. However, these trials have revealed considerable heterogeneity in tumor susceptibility to MV infection. This variability may be linked to differences in MV P gene expression. A single-cell analysis of P gene diversity would inform the development of MV strains with more consistent oncolytic outcomes. In this study, we developed a customized single-cell RNA sequencing (scRNA-seq) platform to capture the MV P gene attenuation site. Using droplet microfluidics, we synthesize barcoding hydrogel beads (BHBs) with primer sequences specific to P gene mRNA. These BHBs were co-encapsulated with single Vero cells infected with attenuated MV. We present data from two independent replicates that demonstrate the use of our platform for generating Illumina sequencing libraries. This study will be further refined to optimize protocols for increased sequencing library yield, enabling multiple sequence alignments and diversity analysis of MV P gene expression across thousands of single infected cells. Application of this platform to the design and validation of improved therapeutic strains will help provide treatments with enhanced oncolytic efficacy.

Introduction

Single-cell sequencing allows for comparison of genomic and transcriptomic profiles from cell to cell. This is especially advantageous for studying highly heterogeneous RNA viruses [1]-[14], such as MV, where diversity in both host cell and virus populations contributes to variable infection outcomes [15]. Single-cell sequencing has been widely applied to the study of RNA viruses, providing insights into the diversity of host and virus gene expression during infection [16]-[17].

Most single-cell sequencing platforms are designed for transcriptomic sequencing of cellular mRNA. These single-cell RNA sequencing (scRNA-seq) technologies were originally developed by the inDrop [18] and Dropseq [19] protocols. scRNA-seq platforms often utilize droplet microfluidics for single-cell isolation and nucleic acid capture [20]. The standard method begins with synthesis of hydrogel beads (HBs) carrying barcoding oligonucleotide primers (BOPs) designed to capture and barcode total mRNA content from a single-cell by targeting the poly-adenylated region of the 3' end. Barcoding hydrogel beads (BHBs) are co-encapsulated with single cells and enzyme mix into microfluidic droplets, where a reverse transcription reaction converts mRNA to cDNA that is barcoded by cell of origin. Drops are then broken, and barcoded cDNA (BC-cDNA) is pooled into a single sample for sequencing library preparation. This method enables high-throughput analysis of mRNA sequences and gene expression diversity across tens of thousands of individual cells.

Standard scRNA-seq methods, that rely on poly-adenylation to capture total mRNA, are unable to resolve specific regions (or genes) of interest and instead only resolve 100-200 base pair regions located near the 3' end of the transcript. While advantageous for surveying the entire

transcriptome, standard scRNA-seq presents a challenge when examining variability in the expression of a single gene. To address this limitation, researchers can generate BHBs carrying custom gene specific primers (GSPs) [21]. These include BHBs designed to target RNA virus genomes [22], and multiplexed HBs capable of capturing both RNA and DNA targets [23]. This customization allows for high-resolution analysis of viral genes involved in specific pathways of interest, such as host immune response.

Measles virus (MV) is a highly pathogenic negative-strand RNA virus that causes severe disease in humans through targeted infection and suppression of host immune systems. During infection, wild type (WT) MVs bind to various immune cells (B lymphocytes, T lymphocytes, dendritic cells, and monocytes) via the CD150 cell surface protein [24]. Additionally, WT MV has been observed to naturally infect cancerous lymphoma cells, which overexpress the CD150 (SLAM) receptor. The MV vaccine strain binds to additional cell types via the CD46 cell surface protein. Many cancer cells overexpress the CD46 receptor, making them more susceptible to MV infection [25]-[28]. This characteristic makes MV a promising candidate for a novel form of oncolytic virotherapy [29]-[32]. Oncolytic virotherapy involves the direct injection of viruses into cancerous tumors, where they selectively replicate in cancer cells, causing syncytia formation and cell death, while also recruiting a localized immune response to the tumor site.

Immunotherapy is widely used in cancer treatment, where specific antigens are introduced to train a patient's own immune system to recognize and eliminate cancerous cells and tumors. Many types of immunotherapies exist, including checkpoint inhibitors, adoptive cell therapy, monoclonal antibodies, immune system modulators, cancer vaccines, and oncolytic virotherapies. MV is highly advantageous for oncolytic virotherapy because it uses a safe and

effective vaccine strain that provides millions of people with lifelong immunity against natural infection each year. MVs are prepared for use as an oncolytic virotherapy by *in vitro* passaging for an extended amount of time in unfavorable conditions, attenuating the virus so that it cannot cause disease but still elicits an immune response. Attenuated MVs carry a mutation in the phosphoprotein (P) gene that results in a tyrosine to histidine shift at position 110 in the amino acid sequence [33]. The P protein plays a critical role in regulating the host immune response to infection [34],[35], with tyrosine 110 required for interferon blocking [34]. Mutations within the P gene, acquired through serial passaging of vaccine strain derivatives, reduce MVs ability to block host interferon responses [33]. These mutations decrease virulence of MV infection and allows for administration of a higher therapeutic dose. However, it also makes the virus more susceptible to clearance by the host immune system. Development of MVs for oncolytic virotherapy therefore requires a balance between attenuation strength, treatment dosage concentrations, and elicitation of the host immune response to tumor sites.

Attenuated MVs can target a wide range of cancerous tumors [32]. However, there remains significant heterogeneity in tumor susceptibility to MV infection during clinical trials [36]-[37]. Image analyses of MV-infected tumor cells reveal cell-to-cell variation in viral replication kinetics and oncolytic phenotypes [38]. Single-cell RNA sequencing of MV-treated tumor cells has identified cell-to-cell heterogeneity in constitutive interferon pathway activation as the primary determinant of MV infection success or resistance [39]. Further single-cell analysis of MV P gene transcripts could determine whether mutations at the attenuation site that emerge during virotherapy treatment is contributing to tumor resistance. This could inform the design of more robust therapeutic strains.

Here, we introduce a new scRNA-seq method to investigate MV gene expression heterogeneity during oncolytic virotherapy treatment. This method highlights the power of integrating microfluidics and scRNA-seq in clinical research, to guide cancer therapy development. We adapted previous protocols [21],[22] to generate custom BHBs designed to capture a 140 nt region of MV P gene mRNA containing the attenuation site required for oncolytic virotherapy. Single infected cells were co-encapsulated with individual BHBs in microfluidic droplets, along with a mixture of enzymes and cell lysis buffer. Barcoded cDNA (BC-cDNA) from millions of individual droplets was pooled into a single sample and processed for Illumina sequencing. The final libraries contained MV P gene transcripts individually barcoded by their drop (i.e., cell) of origin.

This work presents a novel method for exploring MV mutation and gene expression diversity across thousands of single-cells. In doing so, we provide a model system for future investigation of MV diversity during oncolytic virotherapy treatment. This work will enable further description of the genetic differences underlying varied infection phenotypes observed in patients undergoing oncolytic virotherapy treatments and will provide a new method for the design and validation of improved therapeutic strains with enhanced oncolytic efficacy.

Materials and Methods

Cells and Virus

Vero cells were cultured in DMEM supplemented with 10% FBS and grown to confluency. The monolayers were infected with attenuated measles virus (MV) Edmonston B strain derivative expressing green fluorescent protein (MV-GFP) at a multiplicity of infection (MOI) of 2 in tissue culture dishes or 25 cm^2 flasks. MV-GFP was diluted in Opti-MEM media

(Thermo Fischer Scientific) before infection. At 24 hours post infection (hpi) the infected monolayers were treated with 20-40 mM of fusion inhibitory peptide (FIP) (Bachem) to prevent syncytia formation, and re-incubated. Infected cells were collected at 48 hpi by trypsinization, then washed in PBS, and resuspended to $2x10^6$ cells/mL in PBS with 2% BSA.

Microfluidic Device Fabrication

Two microfluidic devices were designed using AutoCAD (Autodesk): (1) a flow focusing device for hydrogel bead synthesis and (2) a three-channel co-flow device for co-encapsulation of hydrogel beads, infected cells, and enzyme/lysis mix. Device designs can be found in the *Supplemental Information*. Microfluidic devices were fabricated in polydimethylsiloxane (PDMS) (Sylgard 184) using standard soft photolithography. Master molds were fabricated by patterning SU-8 photoresist (Kayaku Advanced Materials, SU-8 3050) onto 4 inch silicon wafers (University Wafer, #447) using a micropattern generator (Heidelberg). PDMS was prepared at a 10:1 mass ratio of polymer to cross-linking agent and poured onto the device master molds. Air was purged from the uncured PDMS by placing the filled mold in a vacuum chamber for at least 1 hour. The PDMS was cured at 65 °C for 24 hours and then removed from the mold with a scalpel. Inlet and outlet ports were punched into PDMS slabs with a 1.5 mm biopsy punch (EMS Rapid-Core, Electron Microscopy Sciences). Devices were bonded to 2 x 3 inch glass slides (VWR micro slides, #48382-179) after plasma treatment for 30 seconds at high power and 400 mTorr oxygen pressure, then devices were baked at 80°C for a minimum of 12 hours. The channels of the devices were made hydrophobic by filling the channels with a 0.2 μm filtered solution of 10 μL/mL (tridecafluoro-1,1,2,2-tetrahydrooctyl) trichlorosilane (Gelest) in Novec HFE 7500 fluorinated oil (3M) and incubating for 3 hours at room temperature. Channels were

then flushed with air and devices were baked again for 4 hours at 80°C to evaporate remaining fluorinated oil.

Hydrogel Bead (HB) Synthesis.

30 μm poly-acrylamide hydrogels were synthesized using a flow-focusing microfluidic device. The aqueous phase included 0.1x TBSET buffer, 1x AB solution, 0.30% APS solution, and 400 μM of acyridite-modified DNA Primer 1 (Table 1), prepared in a final volume of 500 μL. To make fluorescent beads 50 μg/mL of 4K MW FITC-Dextran (Sigma-Aldrich, #46944) was suspended in 1x TBSET buffer. The oil phase included 2% PEG-PFPE2-based surfactant (RAN Biotechnologies, #008-FluoroSurfactant) in fluorinated HFE 7500 oil (3M) and 0.4% tetramethylethylenediamine (TEMED) (Thermo Scientific, #17919) in a final volume of 2 mL. See *Supplemental Information* for instructions on how to prepare stock solutions for the aqueous and oil phases. Drop-making was performed using a pressure driven flow system (Fluigent #MFCS-EZ, Le Kremlin-Bicêtre, France), with flow rates of 150 μL/hr for the aqueous phase and 600 μL/hr for the oil phase. Drops were collected into a 5 mL nuclease-free tube loaded with 300 μL of mineral oil. To facilitate hydrogel polymerization, drops were incubated in a water bath at 65°C for 24 hours.

HB Washing

Crosslinked hydrogel beads (HBs) were washed to remove oil and surfactant. All stock solutions for this procedure can be found in *Supplemental Information*. The collection tube began with three phases layered from top to bottom: mineral oil, HBs, and 2% RAN in HFE. The top mineral oil layer and bottom RAN layer were both removed and discarded with a needle and

syringe (BD). HBs were then resuspended in 500 μL of 1x TBSET buffer. To remove remaining RAN surfactant, HBs were washed with 1 mL of 20% 1H,1H,2H,2H-Perfluoro-1-octanol (PFO) in HFE 7500 and centrifuged at 3000 x *g* for 1 minute. This yielded three new phases in the collection tube: TBSET buffer, HBs (milky white / semitransparent), and PFO. The bottom PFO layer was removed and discarded. The PFO wash was repeated until HBs appeared translucent and well-packed. To remove remaining PFO, HBs were washed with 1 mL of 1% Span-80 in hexane and centrifuged at 3000 x *g* for 1 minute. The collection tube then had two phases: Span-80 on top of HBs. The top layer of Span-80 was removed and discarded, and the procedure repeated for a total of 3 washes. To remove residual hexane, HBs were washed in 4 mL of 1x TBSET and centrifuged at 3000 x *g* for 1 minute. The collection tube then had three phases: a thin milky white layer of hexane at the liquid/air interface, TBSET buffer, and a packed translucent mass of HBs at the bottom of the tube. The top two layers of hexane and TBSET were removed and discarded, keeping the pipette tip at the liquid/air interface to ensure removal of hexane. The TBSET wash was repeated twice more, or until the milky white hexane layer was gone. After the final wash, HBs were resuspended in 4 mL of ice-cold 1x TBSET and passed through a 40 μm cell filter attached to a 50 mL nuclease-free tube. The collection tube was washed with an additional 4 mL of ice-cold 1x TBSET and passed through the cell filter. Filtered HBs were centrifuged at 4500 x *g* for 5 minutes to pellet, and the remaining TBSET supernatant was removed, leaving approximately 500 μL of buffer in the tube. HBs can be stored at 4°C for up to 3 months.

<u>BOP Split-and-Pool Ligation</u>

Barcoding oligonucleotide primers (BOP) were ligated onto HBs at acrydite-modified DNA sites (Primer 1) that were incorporated into the poly-acrylamide surface during crosslinking. The procedure required 500 µL of washed and filtered HBs suspended in 1x TBSET buffer. HBs were washed with 4 mL of Hydrogel Bead Wash (HBW) buffer and centrifuged at 3000 x *g* for 1 minute. The HBW supernatant was removed and discarded, and the procedure repeated for a total of 5 washes. The first ligation step (Primer 2, Table 1) added functional sequences for the BclI restriction site, T7 promoter, and Illumina read 1. HBs were resuspended in 3 mL of Primer 2 ligation mix (130 µM of Primer 2, 1x NEB Quick Ligation Buffer, 30 U/µL NEB T7 DNA Ligase, and nuclease free water) and incubated at room temperature for 30 minutes, rotating at 800 rpm. Following incubation, the reaction was centrifuged at 3000 x *g* for 1 minute and supernatant was removed and discarded. HBs were then washed five times with 4 mL HBW, as previously described. After the final wash, 10 µL of HBs in HBW were sampled for downstream ligation QC. The second ligation step (Primer 3, Table 1) added the barcode 1 sequence. HBs were resuspended in 630 µL of Primer 3 ligation mix (1.4x NEB Quick Ligation Buffer and 30 U/µL of NEB T7 DNA Ligase). 22 µL of HB suspension was added to each well of a 96 well plate containing 8 µL of unique Primer 3 sequence per well. Preparation of the Primer 3 to 6 plates (containing barcodes 1 to 4) is described in Table 1. Primer 3 ligation reactions were incubated at 25°C for 30 minutes, rotating at 800 rpm, and inactivated at 65°C for 10 minutes. Individual Primer 3 reactions were pooled by adding 200 µL of ice-cold 1x TBSET buffer to each well and collecting all wells into a single sample. The pooled reaction was centrifuged at 3000 x *g* for 1 minute, and supernatant was removed and discarded. Primer 3 reaction wells were washed a second time with 200 µL of ice-cold 1x

TBSET buffer to collect any remaining HBs. Pooled HBs were washed seven times with 4 mL of ice-cold HBW, as previously described. After the final wash, 10 μL of HBs in HBW were sampled for downstream ligation QC. The Primer 3 ligation procedure was repeated for Primers 4, 5 and 6 (adding barcode 2, 3 and 4 sequences to HBs). Two types of BHBs were synthesized in this work: (1) a single-barcode BHB, with a single barcode across all BHBs, used for protocol validation and control experiments; and (2) a multi-barcode BHB, with different barcode sequences on each bead, which were used for final replicate experiments. The sixth and final ligation step (Primer 7, Table 1) added the functional GSP site. Two GSP were tested, to capture (1) MV P gene mRNA and (2) MV L gene mRNA. The L gene sequence was used as a control and not represented in final replicate experiments. HBs were resuspended in 3 mL of Primer 7 ligation mix (70 μM of primer 7, 1x NEB Quick Ligation Buffer, 30 U/μL NEB T7 DNA Ligase, and nuclease free water) and incubated at 25°C for 1 hour, while rotating at 800 rpm. After incubation, the reaction was centrifuged at 3000 rpm for 1 minute, and supernatant was removed and discarded. HBs were washed five times with 4 mL HBW, as previously described. After the final wash, 10 μL of HBs in HBW were again sampled for downstream ligation QC. To convert the double-stranded DNA ligation product into single-stranded DNA, HBs were washed three times with 4 mL of denaturation buffer. The reaction was inactivated by washing HBs three times with 4 mL of neutralization buffer. At this stage, HBs carried fully ligated single-stranded DNA BOPs. To create an intermediate double-stranded region at the BclI restriction enzyme site and T7 promoter site, HBs were annealed with Primer 8 (Table 1). The reaction was prepared by washing HBs three times with 4 mL of hybridization buffer and resuspending in 400 μL of 1 mM Primer 8. The HB suspension was distributed into PCR tubes, at 30 μL per tube, for

thermocycling. The annealing reaction was incubated at 70°C for 2 minutes then reduced by 3°C per cycle, with ramp rate of 0.1°C/sec, until reaching a final temperature to 25°C. This gave 30 seconds of temperature reduction and 30 seconds of temperature hold, per cycle, which equates to a rate of -3°C/minute. After thermocycling, HBs were pooled by washing each PCR tube with 200 μL of ice-cold 1x TBSET and collecting into a single sample. The PCR tubes were washed a second time to collect any remaining HBs. Pooled HBs were washed five times with 4 mL of 1x TBSET buffer, and filtered through a 40 μm cell filter, as previously described. The HBs, now ligated with BOPs and subsequently referred to as Barcoding Hydrogel Beads (BHBs) were stored at 4°C for up to 3 months. All stock solutions for this procedure can be found in *Supplemental Information*.

Quality Control of Barcoding Hydrogel Beads (BHBs)

The following procedures were used to validate BHB quality prior to single-cell sequencing experiments. All stock solutions are listed in *Supplemental Information*. BHB Size: Image analysis was used to determine the distribution of BHB diameters. BHBs were imaged with a epifluorescence microscope (Nikon Eclipse Ti-U) in the brightfield channel under phase contrast, at 10x and 20x magnification with an additional ROI zoom. The camera was set to 10 DIA (disascopic/brightfield illumination), 12-bit depth, and 1 ms exposure for both magnifications. BHBs synthesized with FITC-dextran were imaged in the FITC channel under phase contrast at 10x and 20x magnification with an additional ROI zoom. The camera was set to 10 DIA and 12-bit depth for all magnifications, with a 100 ms exposure for the 10x images 300 ms exposure for the 20x images. BHB diameter was measured using the ‘object count’ analysis in the Nikon NIS-Elements software. BOP Length: BHBs sampled after each BOP ligation step

were treated with a restriction enzyme digest to release primers from beads for quantification of BOP length. Restriction digest reactions were prepared at a final volume of 20 μL (2 μL BHBs, 2 μL of 10x NEB r3.1 Buffer, 1 μL of 10 U/μL NEB BclI restriction endonuclease, 15 μL of nuclease-free water) for each BOP ligation sample, in duplicate. Reactions were incubated at 50ºC for 30 minutes, while rotating at 800 rpm. Capillary electrophoresis (Agilent, TapeStation D1000 kit) was used to determine BOP length after each ligation step. The expected size of a full-length product was 218 nucleotides (nt). This procedure also validated function of the restriction enzyme site. Fluorescent in situ hybridization (FISH) Assay: To further validate the addition of exact BOP sequences, BHBs were incubated with Cy5 FISH probes. The sequence of each probe (Table 2) was complementary to a different sequence on each of the individually ligated primers (2 to 7). For each individual FISH assay, 10 μL of BHBs were resuspended in 1.2 mL of hybridization buffer and centrifuged at 1000 x *g* for 30 seconds. 1,164 μL of supernatant was removed and discarded, leaving 36 μL of suspended BHBs. 4 μL of 100 μM of a single FISH probe was added to a single sample of BHBs. The number of samples tested varied by BHB type, 6 probes for the single-barcode BHBs and 2 probes for the multi-barcode BHBs. FISH reactions were incubated at room temperature for 20 minutes away from light. BHBs were then washed three times with 1.2 mL of hybridization buffer. After the final wash, 1,164 μL of supernatant was removed and discarded, again leaving 36 μL of suspended BHBs remaining for image analysis. BHBs were imaged in the Cy5 channel under phase contrast at 10x and 20x magnification with an additional ROI zoom. The camera was set to 10 DIA and 12-bit depth for all magnifications, with a 100 ms exposure for the 10x images 300 ms exposure for the 20x images. cDNA Synthesis and Barcoding in Bulk: BHB function was validated using a bulk assay

in well plates. Reverse transcription reactions were prepared at 20 μL per reaction (6.4 μL of 60% Optiprep, 4 μL of 5X FSS Buffer, 1 μL of 10% Trition-x100, 1 μL of 0.1 mM DTT, 1 μL of 10 mM dNTP, 1 μL of Rnase-Away nuclease inhibitor, 1μL BclI of 10,000 U/mL restriction endonuclease, 1 μL of 200 U/μL Supercript III Reverse Transcriptase, 1 μL of BHBs, 2.6 μL of RNA template). Two RNA templates were tested, (1) a synthetic RNA gene block at 10^9 copies/μL (Table 3) and (2) MV infected Vero cells. Each bulk cDNA synthesis experiment included 24 individual 20 μL reactions (~500 μL pooled reaction volume), to equal the volume expected from droplet cDNA synthesis. Reverse transcription reactions were incubated at 50°C for 1 hour, inactivated at 70°C for 15 minutes, and cooled on ice for 1 minute. Before measuring cDNA yield, the product was purified using the procedure listed in *cDNA Synthesis and Barcoding in Drops*.

Co-Encapsulation of BHBs and MV Infected Cells

ScRNA-seq experiments began with 500 μL of BHBs suspended in TBSET buffer. BHBs were washed three times in 4 mL of pre-encapsulation buffer (0.6% Triton-X100, 150 mM Tris-HCl pH 7.4, and nuclease-free water). BHBs were then resuspended in 200 μL of activation buffer (0.6% Triton-X100, 150 mM Tris-HCl pH 7.4, 15 mM DTT, 2.5x NEB FSS Buffer, and nuclease-free water) and incubated at room temperature for 10 minutes. Activated BHBs were centrifuged at 4500 *x g* for 30 seconds, the supernatant was discarded, and BHBs were stored on ice until device loading. MV infected Vero cells were resuspended in 60% Optiprep (Sigma-Aldrich, #D1556) at $1x10^6$ cells/mL and stored on ice until device loading. The enzyme and cell lysis mixture was prepared at 500 μL final volume (1.25x NEB FSS Buffer, 5mM DTT, 1.5 mM

dNTP, 0.9% Triton-X100, 25 U/μL SuperScript III Reverse Transcriptase, 1.07 U/μL NEB BclI Restriction Endonuclease, 2.5 U/μL SUPERaseln nuclease inhibitor, and nuclease-free water) and stored on ice for the remainder of the procedure, including during drop-making. 3 mL of 3% RAN-008 surfactant in HFE 7500 was prepared for drop-making. The stock solutions used in this procedure are listed in *Supplemental Information*. To prepare the drop-making setup, BHBs were loaded into a 1 mL syringe backed with HFE 7500. Infected cells were transferred to a 2 mL tube containing a stir bar and set up on a magnetic stir plate. The enzyme/lysis mix was transferred into a 2 mL tube and stored on ice. The oil phase was transferred into a 2 mL tube and set at room temperature for drop-making. A 3-point co-flow device [22] was used to encapsulate BHBs and MV infected cells with the enzyme/lysis mix required for cDNA synthesis and barcoding. Cells, enzyme mix, and oil were injected into the device using a pressure driven flow system (Fluigent #MFCS-EZ, Le Kremlin-Bicêtre, France). BHBs were injected into the device using a syringe pump (New Era, NE-1000). The co-encapsulation device was set up on a Nikon Eclipse Ts2 mounted with a high-speed camera (Phantom) to visualize drop-making. To begin drop-making, the oil phase was injected first, followed by the cell phase, the enzyme phase, and the BHB phase. Drop-making was conducted at the following pressures/flowrates: oil phase (120 psi), MV infected cells (100 psi), enzyme/lysis mix (90 psi), BHBs (35 μL/hr). 500 μL of drops were collected into a 2 mL nuclease-free tube on ice.

cDNA Synthesis and Barcoding in Drops

cDNA Synthesis: Encapsulated cells and BHBs were incubated in a water bath at 50°C for 1 hour to initiate cDNA synthesis and barcoding. During incubation, the BclI restriction enzyme functioned to release BOPs from BHBs. At the same time, the lysis mix acted to break down cell and virus membranes and release MV mRNA. The GSP then primed reverse transcription of mRNA to cDNA that is tagged with the unique barcode sequence carried on each individual BHB. This reaction was inactivated by incubating drops at 70°C for 15 minutes in a heat block, then on ice for 1 minute. Droplet emulsions were then broken by adding 1 mL of 10% Perfluoro-octanol (PFO) in HFE 7500 to drops, vortexing to mix, and centrifuging at 5000 *x g* for 30 seconds to phase separate. The PFO washing procedure was repeated until all drops were broken. The bottom oil layer was removed and discarded, while BC-cDNA from the coalesced aqueous phase was pooled and collected into a new tube. The pooled BC-cDNA was purified with a 3-step procedure**. Column filtration cleanup:** To remove BHBs from BC-cDNA, the pooled sampled was added to 0.45 μm filter column and centrifuged at 12,000 *x g* for 2 minutes. The filter was rinsed with 10 μL of nuclease-free water and centrifuged once more. cDNA filtrate was kept and the filter discarded. Exonuclease cleanup: To remove unused ssDNA BOPs, BC-cDNA was treated with a restriction endonuclease in 100 μL reactions (60 μL of cDNA, 30 μL of 10x R3.1 buffer, 10 μL of 20 U/μL ExoI restriction endonuclease) in PCR tubes incubated at 37°C for 1 hour. The reaction was inactivated at 80°C for 5 minutes. Magnetic bead size selection cleanup: BC-cDNA was size selected by combining with 1x volume of AMPure XP magnetic beads (Beckman Coulter). The mixture was vortexed for 30 seconds and spun to combine. The reaction was incubated at room temperature for 10 minutes, then placed tube on a magnet holder for 15 minutes or until the sample liquid appeared clear. The supernatant was

carefully removed and discarded, being sure to not disturb or remove the magnetic bead pellet. Keeping the tube on the magnet, 1 mL of 80% EtOH was added to wash the magnetic beads. Beads were incubated with EtOH for 2 minutes at room temperature. EtOH was then removed and discarded. This EtOH washing procedure was repeated twice more. After the final EtOH wash, all EtOH was removed from the tube and, with the tube lid open, beads were air dried for 2 to 3 minutes (until the bead pellet began to turn from a dark wet brown to a light brown around the edges). Once dry, the sample tube was removed from magnetic holder and cDNA was eluted into nuclease-free water at a volume equal to the starting volume. The elution was incubated at room temperature for 10 minutes and then transferred back to the magnet holder for 15 minutes or until the sample liquid appeared clear. The cDNA elution was collected and transferred to a new tube. qPCR for cDNA quantification: BC-cDNA product was verified with a SBYR-based qPCR assay. Custom primers were designed to capture a spacer region between the barcode 4 and UMI sequences (forward primer) and the 3' end of the MV P gene mRNA region of interest (reverse primer). A schematic is shown in Figure 15A, and custom primer sequences are listed in Table 4. Each qPCR reaction was prepared at a 20 µL final volume (2 µL of cDNA cleanup, 10 µL of 2X SYBR Green PCR Master Mix, 0.5 µL of 10 µM Forward Primer, 0.5 µL of 10 µM Reverse Primer, 7 µL of nuclease-free water). qPCR reactions were incubated in a QuantStudio 3 machine at the following thermocycling conditions: 50°C for 2 minutes (1x); 95°C for 10 minutes (1x); 55°C for 30 seconds, 72°C for 30 seconds (45x). A melt curve was run immediately following thermocycling, with the following conditions: 95°C for 15 seconds, 60°C for 1 minute, and 95°C for 1 seconds. The heated cover was set to 105°C and all ramp rates were set to 1.6 °C /second. Amplification and melt curves were analyzed using the QuantStudio

Design and Analysis software. TapeStation for cDNA length: The length of the BC-cDNA qPCR product was measured using Capillary Electrophoresis (Agilent, TapeStation D1000 kit).

Sequencing Library Preparation

The following procedures were used to prepare BC-cDNA for Illumina NovaSeq 150 bp paired-end sequencing. Second Strand Synthesis (SSS): BC-cDNA was converted to double stranded DNA using the NEB Next Ultra II Non-Directional RNA Second Strand Synthesis Kit. On ice, BC-cDNA was separated into 20 µL reactions (15µL of cDNA, 2µL of NEB 10x SSS buffer, 1 µL of NEB SSS enzyme, and 2 µL of nuclease-free water) and incubated at 16°C for 2.5 hours. After incubation, all SSS reactions were pooled to a single sample. The SSS DNA was purified using 1.2x AMPure XP magnetic beads, using the magnetic bead size selection cleanup previously described. SSS DNA was eluted in nuclease-free water at a final volume equal to 8 µL per SSS reaction. In-Vitro Transcription (IVT): Purified SSS DNA was converted to IVT RNA using the NEB HiScribe T7 High Yield RNA Synthesis Kit. A NTP/buffer master mix was prepared by combining equal volumes of 10x T7 reaction buffer, ATP, GTP, UTP, and CTP. SSS DNA was separated into 20 µL reactions (8 µL of SSS DNA, 10 µL of NTP/buffer master mix, and 2 µL of T7 polymerase) and incubated overnight at 37°C for 12 hours. The IVT RNA can be stored at 4°C for up to 2 hours after the reaction is complete. Individual IVT reactions were pooled to a single sample and purified with 1.3x AMPure XP magnetic beads, using the magnetic bead size selection cleanup previously described. IVT RNA was eluted in nuclease-free water at a final volume of 20 µL per IVT reactions. The concentration of IVT RNA was then quantified using the Qubit ssRNA kit. Secondary Reverse Transcription (RT2): To capture a region of interest on the MV mRNA, a second gene specific primer (Primer 9, Table 5) was used

for a second reverse transcription reaction on IVT RNA. This step also allowed for addition of the Illumina read 2 sequence that is required for sequencing. IVT RNA was separated into 20 µL reactions (12 µL of IVT RNA, 1 µL of 100 µM Primer 9, 1 µL of 20 U/µL SUPERaseln, 1 µL of 10 mM dNTP, 1µL of 200 U/µL SSIII Enzyme, 4 µL of 5x FSS Buffer) and incubated at 55°C for 1 hour. The reaction was inactivated at 70°C for 15 minutes. Individual RT2 reactions were pooled to a single sample and purified with 1.2x AMPure XP magnetic beads. Purified cDNA2 was eluted in nuclease-free water at a final volume equal to 20 µL per RT2 reaction. PCR amplification and Illumina Adaptor Ligation: cDNA2 was PCR amplified prior to sequencing. The PCR amplification also added Illumina P5 and P7 adaptor sequences via Primer 10 and Primer 11 (Table 5). cDNA2 was separated into 50 µL PCR reactions (5 µL of RT2 cDNA, 1x GC buffer, 0.3 mM dNTPs, 0.5 µM of Primer 10, 0.5 µM of Primer 11, 0.5 U Roche KAPA HiFi HotStart Polymerase) and incubated at the following thermocycling conditions: 98°C for 3 minutes (1x); 98°C for 20 seconds, 55°C for 20 seconds, 72°C for 30 seconds (3x); 98°C for 20 seconds, 65°C for 20 seconds, 72°C for 30 seconds (24x); 72°C for 30 seconds (1x). PCR amplified libraries were purified with 0.7x AMPure XP magnetic beads (according to previously described protocols) and eluted in nuclease-free water at a final volume of 20 µL per PCR reaction.

Quality Control of Sequencing Library Preparation

qPCR for Illumina sequencing library quantification: Illumina P5/7 adaptor ligation was confirmed with the Roche KAPA Library Quantification Kit. Libraries were diluted 10-fold in series from 10^{-1} to 10^{-5} in nuclease-free water. A master mix of dye and primer (1 mL of 2x KAPA SYBR FAST qPCR Master Mix, 40 µL of 50x ROX HIGH, and 200 µL of 10x Primer

Premix) was combined with each library dilution in a 20 μL qPCR reaction (12.4 μL of DYE/Primer master mix, 4 μL of PCR product dilution, 7.2 μL of nuclease free water). Reactions were incubated with the following thermocycling conditions: 95°C for 5 minutes (1x); 95°C for 30 seconds, 60°C for 45 seconds (35x); Melt 65°C to 95°C. Amplification and melt curves were analyzed using the QuantStudio Design and Analysis software. Qubit for PCR product concentration: library concentration was determined with the Qubit dsDNA Broad Range Assay kit, completed according to manufacturer's instructions. TapeStation and Bioanalyzer for library nt length: Average library length was measured using capillary electrophoresis (Agilent, TapeStation HSD1000 kit) and (Agilent, Bioanalyzer High Sensitivity kit).

Sequence Assembly

Raw sequencing data were processed following the Drop-seq Core Computational Protocol v2.0.0 developed by James Nemesh in Steve McCarroll's lab at Harvard Medical School. The analysis was conducted using Drop-seq tools (v3.0.2), Picard (v3.3.0), and STAR (v2.7.11b) by Ben Hur, PhD. In brief, the paired-end FASTQ files were first converted into an unaligned BAM format using Picard 'FastqToSam', preserving read pairing information. Cell barcodes and unique molecular identifiers (UMIs) were extracted from barcode reads (read 1) using 'TagBamWithReadSequenceExtended'. The first 80 bases of barcode reads (BASE_RANGE=1-80) were assigned as the cell barcode (CB) tag, while bases 81–89 (BASE_RANGE=81-89) were designated as the UMI (CM) tag. Only bases with a Phred quality score ≥10 (BASE_QUALITY=10) were retained, except when at most one base (NUM_BASES_BELOW_QUALITY=1) fell below this threshold. Reads lacking valid cell

barcodes or UMIs were then filtered using ‘FilterBam’, ensuring that only high-confidence reads were retained for downstream analysis. ‘TrimStartingSequence’ was used to remove adapter sequences for the barcode reads that may introduced during library preparation (SEQUENCE=ACACTCTTTCCCTACACGACG), allowing no mismatches (MISMATCHES=0). Additionally, up to five bases (NUM_BASES=5) were removed as needed to eliminate residual adapter sequences. Preprocessed BAM files were converted to FASTQ format using Picard ‘SamToFastq’. The resulting FASTQ files were then aligned to the custom attenuated MV genome (16,656 base pairs) using with STAR default parameters.

Results

Single-cell mRNA Sequencing of Attenuated Measles Virus (MV) Using Droplet Microfluidics.

The workflow began with the production of BHBs designed to capture a specific region of MV P gene mRNA (Fig 5A). The BHBs were synthesized as poly-acrylamide hydrogels functionalized with acrydite-modified dsDNA. Six BOPs were sequentially ligated to the BHBs, incorporating sequences for mRNA capture, cDNA synthesis, droplet-based barcoding, and sequencing library preparation. These BHBs were co-encapsulated with individual vero cells infected with a MV vaccine strain derivative, currently used in clinical trials for oncolytic virotherapy, at an MOI of 2 for 48 hours post infection (hpi) (Fig 5B). BHBs and Cells were encapsulated into microfluidic drops along with a mixture of enzymes and cell lysis buffer used for cDNA synthesis and barcoding and incubated at 50°C for one hour to initiate the reaction (Fig 5C). MV P gene mRNA was barcoded by cell of origin and prepared for Illumina sequencing. Sequencing reads were aligned to an MV reference genome (Fig 5D) to show proof

of concept, and future work will compare sequence diversity across drops using multiple sequence alignments and principal component analysis (Fig 5D).

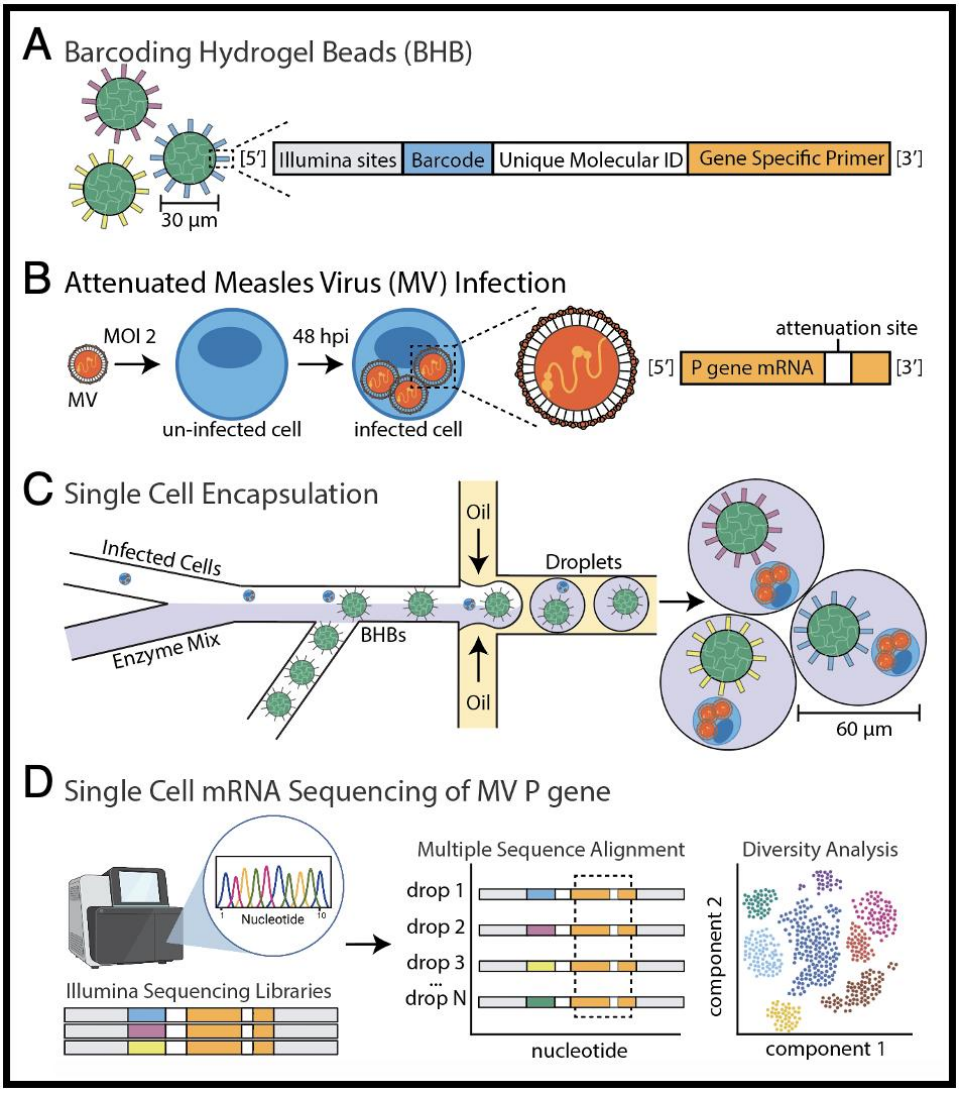


Figure 5. Overview of Single-cell mRNA Sequencing of Attenuated Measles Virus (MV) Using Droplet Microfluidics. (A) 30 µm barcoding hydrogel beads (BHBs) were designed to carry dsDNA primers with sequences for Illumina library preparation, a barcode that is unique to each bead, a molecular ID that is unique to each primer, and gene specific primer (GSP) that varies by mRNA target region. (B) BHBs were synthesized with GSPs designed to capture a 140 nt region of MV P gene mRNA containing the attenuation site required for oncolytic virotherapy. (C) A microfluidic device was used to co-encapsulate BHBs with MV infected Vero cells harvested at 48 hours post infection (hpi). Single BHBs and single-cells were combined in 60 µm drops with a mixture of enzymes and cell lysis reagents required for cDNA synthesis and barcoding of P gene mRNA. Drops were broken and cDNA pooled for sequencing library preparation. (D) All libraries were sequenced with Illumina NovaSeq 150-bp paired end sequencing. Consensus

sequences were assembled by alignment to a reference genome. Future work will use sequencing alignments to compare attenuation site diversity across single-cell infections.

Barcoding Hydrogel Bead (BHB) Synthesis

Each BHB consisted of an HB adapted with BOPs to include sequences (Fig 10) for a BclI restriction site, T7 promoter site, Illumina read 1 primer, one unique barcode per bead, one unique molecular identifier (UMI) per primer, and a conserved GSP (Fig 6A). The GSP used in this work was designed to target a 140 nt region of MV P gene mRNA containing the attenuation site used for oncolytic virotherapy. HB synthesis required a flow-focusing microfluidic device with 40 μm channel heights (Fig 6A, i). The HBs were made of polyacrylamide, with polymerization initiated by TEMED in the oil phase interacting with APS in the aqueous phase to form APS free radicals that facilitated the crosslinking of acrylamide and bis-acrylamide monomers. HBs were initially synthesized to contain FITC-dextran fluorescent dye for better visualization during device testing experiments (SI Fig 6A and 5.2B). After synthesis, washing, and filtration (SI Fig 7A), HBs measured an average of 33.5 μm ± 2.1 μm in diameter, with a coefficient of variation *(CV)* of 6.1% (SI Fig 7B).

Functional BOPs were assembled by ligation of seven individual primers (Table 1), with complementary overhanging regions on both the 3’ and 5’ ends, and phosphorylation on the forward strand 5’ end. Sequential BOP ligation (Fig 6A, iii) was facilitated by T7 DNA ligase catalyzing phosphodiester bonds between the complementary overhanging regions. Primer 1 was incorporated into the BHB surface during cross-linking and served as a template for the addition of Primer 2 (containing the BclI restriction site, T7 promoter, and Illumina read 1 sequence) to all BHBs. Ligation of Primers 3 to 6 was more complex, requiring split-and-pool ligation to facilitate the addition of four unique barcode sequences to each BHB. This was done according

to previously published protocols [21][22] by splitting BHBs into each well of a 96-well plate, containing a single unique barcode sequence, conducting the ligation reaction, then pooling into a single sample after each ligation step. The split-and-pool method allows for production of up to 96^4 (8.4×10^7) unique barcodes in a 500 µl batch of BHBs. Control BHBs were also produced to carry a single known barcode sequence, for use in protocol validation and QC experiments (Fig 11). Following the ligation of unique barcodes, all BHBs were ligated with Primer 7 to add the GSP sequence for capturing MV P gene mRNA (SI Fig 5A). Primer 7 also included a unique molecular identifier (UMI) sequence, randomly generated during synthesis, that functioned as a PCR amplification bias control by individually tagging each read with a unique sequence that should be evenly distributed within the final library. Overexpression of a single UMI sequence would indicate PCR amplification bias. BHBs were sampled after each ligation step to measure intermediate nt length, shown by the top band of each column in Fig 6B. This QC step tests for successful ligation across all primers. The expected length of the final BOP ligation product was 218 nt (SI Fig 5B), and across three biological replicates we measured BOP lengths of 255 nt, 268 nt, and 248 nt (Fig 6B, red box) via the TapeStation capillary electrophoresis assay. This was in alignment with single-barcode BHBs (SI Fig 2C and 2D), whose average BOP lengths measured 253 nt and 261 nt.

Following ligation of Primer 2 to Primer 7, further modification of BOP structure was required to make BHBs functional for cDNA synthesis and barcoding. The first modification was chemical denaturing of the dsDNA structure required for ligation to ssDNA necessary for reverse transcription. Once denatured, an additional ssDNA Primer 8 (Table 1) was annealed to the BOP

region containing BclI and T7 promoter sequences. This is because the enzymes that recognize the BclI and T7 promoter sites bind to dsDNA.

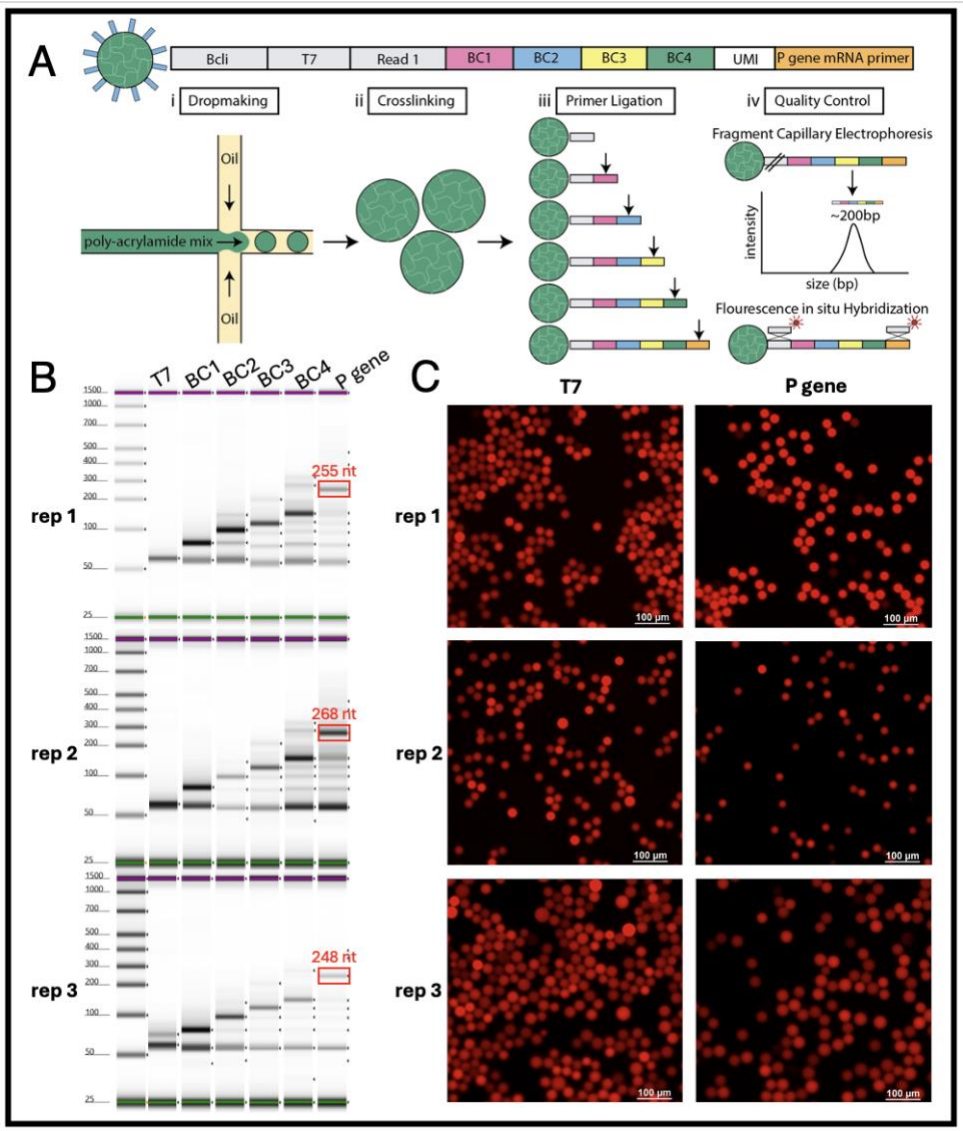


Figure 6. Barcoding Hydrogel Bead (BHB) Synthesis. (A) Polyacrylamide hydrogel BHBs were synthesized using a flow-focusing microfluidic device (i). BHBs were made with acyrdite-modified DNA oligonucleotides that incorporated within multiple sites on the bead surface during crosslinking (ii). These DNA oligos served as anchor points for the ligation of functional primer sequences (iii), including a BclI restriction enzyme site, T7 promoter site, Illumina read 1 sequence, 4 unique barcode sequences per bead, and a gene specific primer complementary to MV P gene mRNA. The BHB ligation was validated (iv) with fragment capillary electrophoresis and fluorescence in situ hybridization (FISH) assays. (B) Ligation intermediates showed increasing nt length (top band of each column) after each step, and final BOP product lengths

(255 nt, 268 nt, and 248 nt) close to the expected length of 218 nt. (C) Final BOP sequences were visualized with Cy5 FISH probes for the T7 promoter site and MV P gene mRNA primer site.

Following final modification, BHBs were visualized using fluorescence in situ hybridization (FISH) (Fig 6C). Small (~20nt) DNA probes carrying a Cy5 fluorophore were designed with sequences complementary to the T7 promoter sequence on Primer 2 and the P gene GSP on Primer 7 (Table 2). Additional FISH probes were designed for control BHBs carrying a single barcode sequence (Table 2), to validate the ligation of Primers 3 to 6 (SI Fig 6A and 5.2B). Samples of complete BHBs were separately incubated with each FISH probe, washed, and imaged under fluorescence microscopy. FISH fluorescence imaging showed that both T7 (Primer 2) and P gene GSP (Primer 7) sequences were present on the final BHB products across three biological replicates (Fig 6C).

Droplet cDNA Synthesis and Barcoding

BHBs were co-encapsulated with single MV infected cells and enzyme/lysis mix using a droplet microfluidic platform (Fig 7A). Drop-making was conducted with a previously designed microfluidic device [22] (Fig 7B). Device operation was optimized for loading single BHBs with single-cells. To accomplish single BHB loading, we examined the effect of flow rate (μl/minute) fraction (Q_{BHB}/Q_{oil}) on BHB loading, testing three conditions: $Q_{BHB}/Q_{oil} = 0.16$ (4 μl/minute BHB and 25 μl/minute oil), $Q_{BHB}/Q_{oil} = 0.20$ (6 μl/minute BHB and 30 μl/minute oil), $Q_{BHB}/Q_{oil} = 0.24$ (6 μl/minute BHB and 25 μl/minute oil) (Fig 7C). Image analysis was used to count the number of BHBs per drop from ~100 drops per condition (Fig 13). The flow rate fraction that produced the highest frequency of single BHBs in drops (86%) and lowest frequency of empty

drops (14%) was $Q_{BHB}/Q_{oil} = 0.24$ (Fig 7C, triangle points with dotted line) and therefore used in subsequent BHB encapsulation experiments.

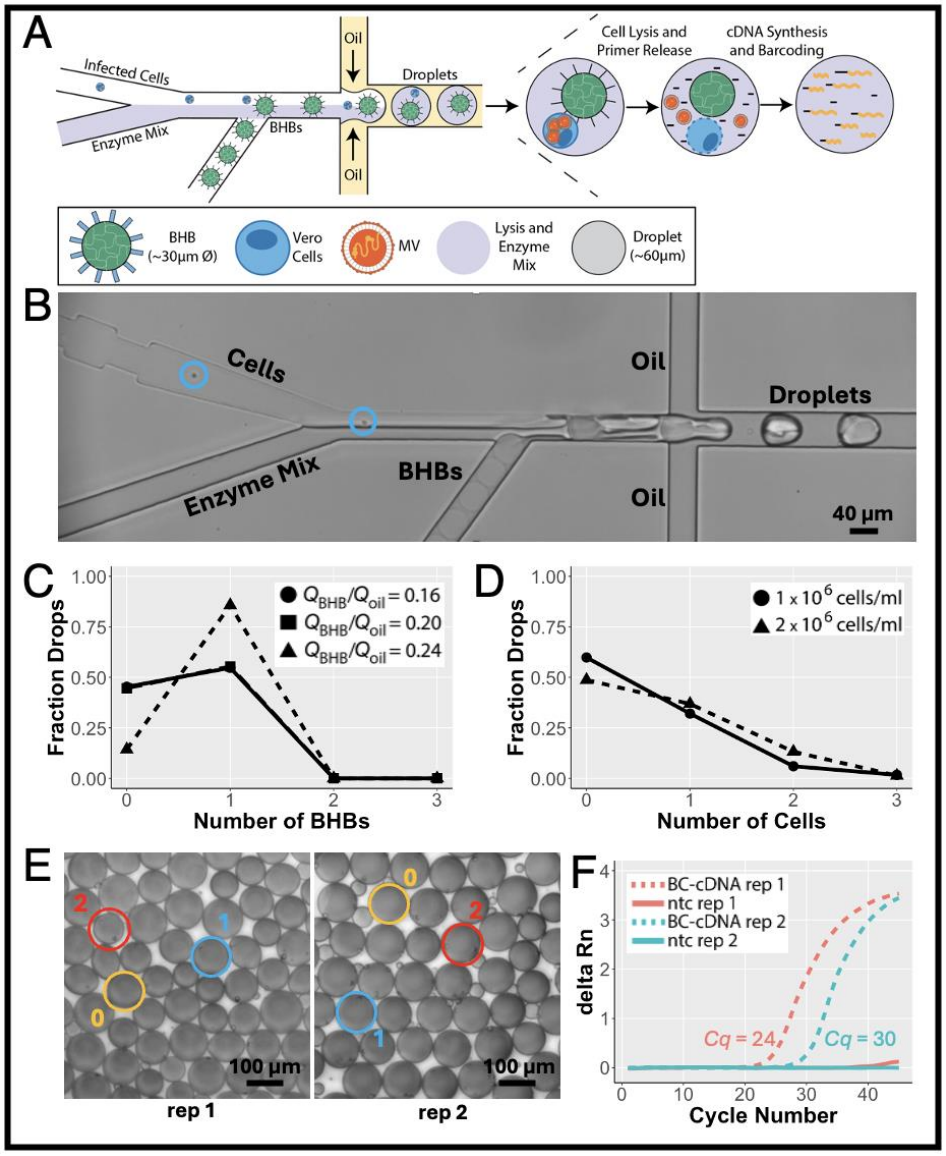


Figure 7. Droplet cDNA Synthesis and Barcoding. (A) BHBs and MV infected Vero cells were co-encapsulated with lysis/enzyme mix into 60 µm droplets. Following encapsulation, drops were incubated at 50°C for 1 hour to allow for cell and virus lysis, digestion of BOPs from BHBs by the BclI restriction enzyme, and reverse transcription of MV P gene mRNA into cDNA barcoded by cell of origin (B) 3-channel co-flow device [22] was optimization for loading of single BHBs with single-cells. (C) Device optimization of BHB and oil flow rate fractions (Q_{BHB}/Q_{oil}), counting the number of BHBs per drop. (D) Further optimization of cell suspension concentration, counting the number of cells per drop. (E) Optimized device operation conditions were used in two biological replicate scRNA-seq experiments. Representative images show

drops containing either 0 (yellow circle), 1 (blue circle), or 2 (red circle) cells per drop. (F) Barcoded cDNA (BC-cDNA) was detected from each of the two scRNA-seq replicates using a custom qPCR assay that amplifies a region covering both the BOP and the MV P gene cDNA sequences.

Drop-making was further optimized for single-cell loading by encapsulating two suspension concentrations, $2x10^6$ cells/mL and $1x10^6$ cells/mL (Fig 7D), with BHBs and enzyme mix using the flow rate fraction determined in Fig 7C. Following encapsulation, droplets were imaged and the number of cells per drop counted from ~150 drops for the $2x10^6$ cells/mL condition and ~230 drops for the $1x10^6$ cells/mL condition (Fig 14A). Both suspensions showed similar single-cell loading (37% for $2x10^6$ cells/mL and 32% for $1x10^6$ cells/mL) but $1x10^6$ cells/mL had a lower rate (5%) of two cells per drop than $2x10^6$ cells/mL (13%) and was therefore used for remaining experiments. Representative images of cell loading for two independent scRNA-seq experiments are shown (Fig 7E).

After co-encapsulation, drops were incubated at 50°C for 1 hour to allow infected cells and virus particles to lyse, the Bc1I enzyme to digest BOPs from BHBs, and the reverse transcriptase enzyme to facilitate cDNA synthesis and barcoding of MV P gene mRNA. After incubation, drops were broken and the BC-cDNA pooled into a single sample for purification. Pooled and purified BC-cDNA was detected using qPCR, with custom primers (Table 4) to amplify a portion of the BOP sequence and the full P gene region of interest (Fig 15A). Two independent scRNA-seq experiments detected BC-cDNA (SI Fig 6B) at cycles $Cq = 24$ and $Cq = 30$ (Fig 7F).

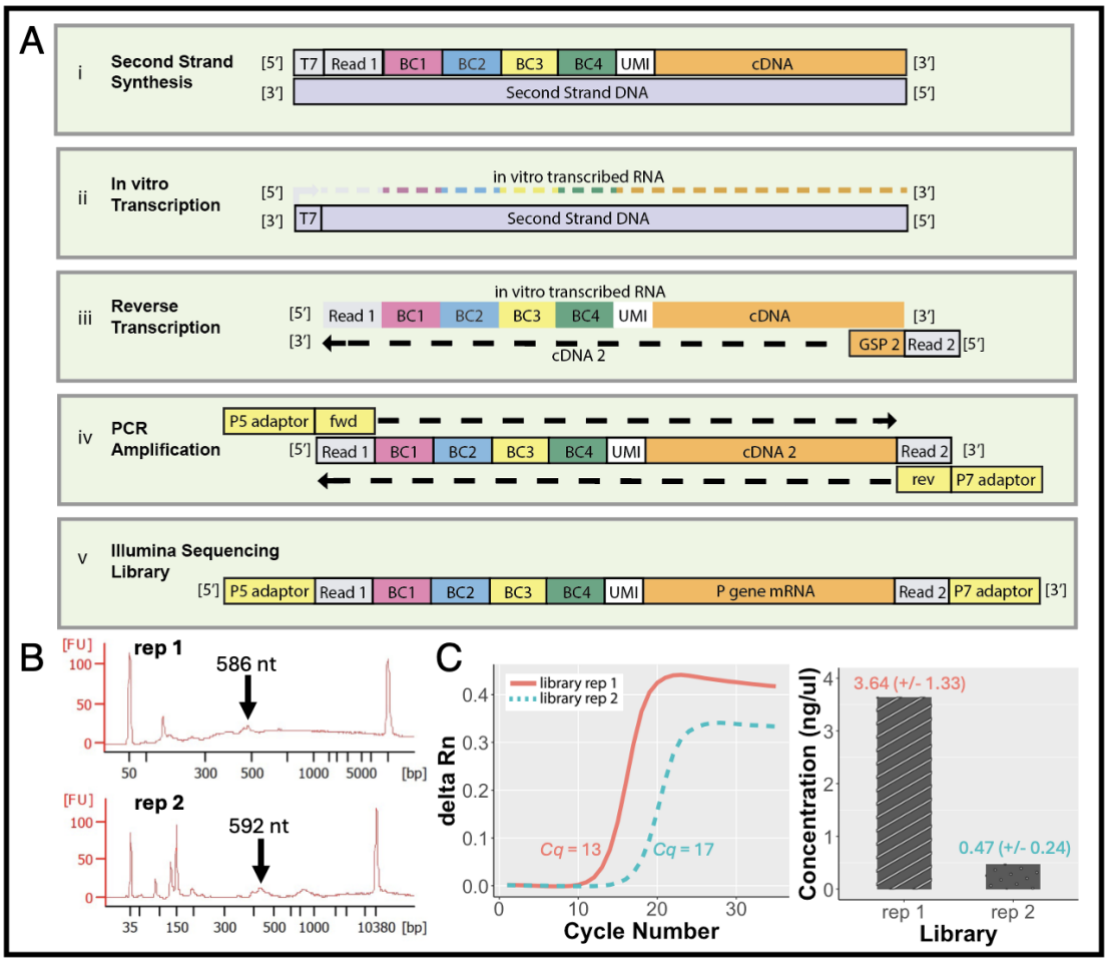


Figure 8. Illumina Sequencing Library Preparation. (A) Illumina sequencing libraries were prepared using the following method (i) BC-cDNA was first converted to dsDNA through a second strand synthesis reaction. (ii) This dsDNA was the template for an *in vitro* transcription reaction, to convert the product to RNA. (iii) IVT RNA was then reverse transcribed using a second gene specific primer designed to capture a specific 140 nt region of interest on MV P gene mRNA. The cDNA 2 product was also adapted with the Illumina read 2 site at this step. (iv) Libraries were then amplified and adapted with the Illumina P5 and P7 sites in a PCR step. (v) The final libraries should include the Illumina P5 and P7 adaptor sequences, Illumina read 1 and 2 sequencing primer sites, an 80 nt barcode sequence, a 9 nt UMI sequence, and a 142 nt region of cDNA derived from the MV P gene mRNA attenuation site. (B) scRNA-seq libraries were prepared for two independent biological replicates. The length of each Illumina sequencing library was measured by capillary electrophoresis, using the Bioanlyzer D1000 kit (rep 1) and HSD1000 kit (rep 2). (C) Each library was detected and quantified by qPCR, using the KAPPA Illumina Library Quantification Kit.

Sequencing Library Preparation

BC-cDNA collected from ~4 million drops containing ~1.28 million cells was prepared for Illumina sequencing using a non-standard library preparation procedure [21][22] (Fig 4A). Here, cDNA was first converted into dsDNA with a second strand synthesis reaction (Fig 8A, i). The dsDNA intermediate was *in vitro* transcribed (IVT), using the T7 protomer sequence previously ligated onto BHBs, into antisense RNA (Fig 8A, ii). The IVT reaction served to also amplify the RNA, as individual strands of barcoded DNA were transcribed multiple times. IVT RNA was then reverse transcribed using a second gene specific primer (GSP) (SI Table 5.5, Primer 9) to obtain BC-cDNA that contained sequence from a specific 140-bp region of the MV P gene mRNA (Fig 8A, iii). The secondary GSP used at this step also carried the Illumina read 2 primer site, required for downstream sequencing. The second pool of cDNA was PCR-amplified, using primers that target the Illumina read 1 and 2 sites found on each library and that also carried sequences for the Illumina P5 and P7 adaptors (SI Table 5.5, Primers 10 and 11) to be added to each library during PCR (Fig 8A, iv). Libraries had final lengths (Fig 8B) and concentrations (Fig 8C)(Table 6) averaging of 586nt at 3.6 ng/μL (rep 1) and 592 nt at 0.47 ng/μL (rep 2).

Assembly and Analysis of MV P Gene mRNA scRNA-seq

Single-cell P gene mRNA sequences were assembled by aligning reads to a reference genome for the MV vaccine strain derivative used in clinical trials for oncolytic virotherapy [39]. Reads were first trimmed of adaptor and Illumina primer sequences. Then high quality reads (Q score >20) were mapped to the reference genome. Sequence alignments were visualized using integrative genome viewer (IGV) (Figure 9A) and showed high genome coverage across a 160 nt

region of MV P gene. While the reads that mapped to the reference genome did so at high coverage, most reads from the two scRNA-seq replicate experiments did not map. This may be due to low read 2 quality (Figure 9B), where read 1 contains the barcode sequence and read 2 contains the targeted P gene mRNA sequence. Future optimization of the scRNA-seq protocol should focus on increasing product yield to improve sequencing results for population diversity analysis.

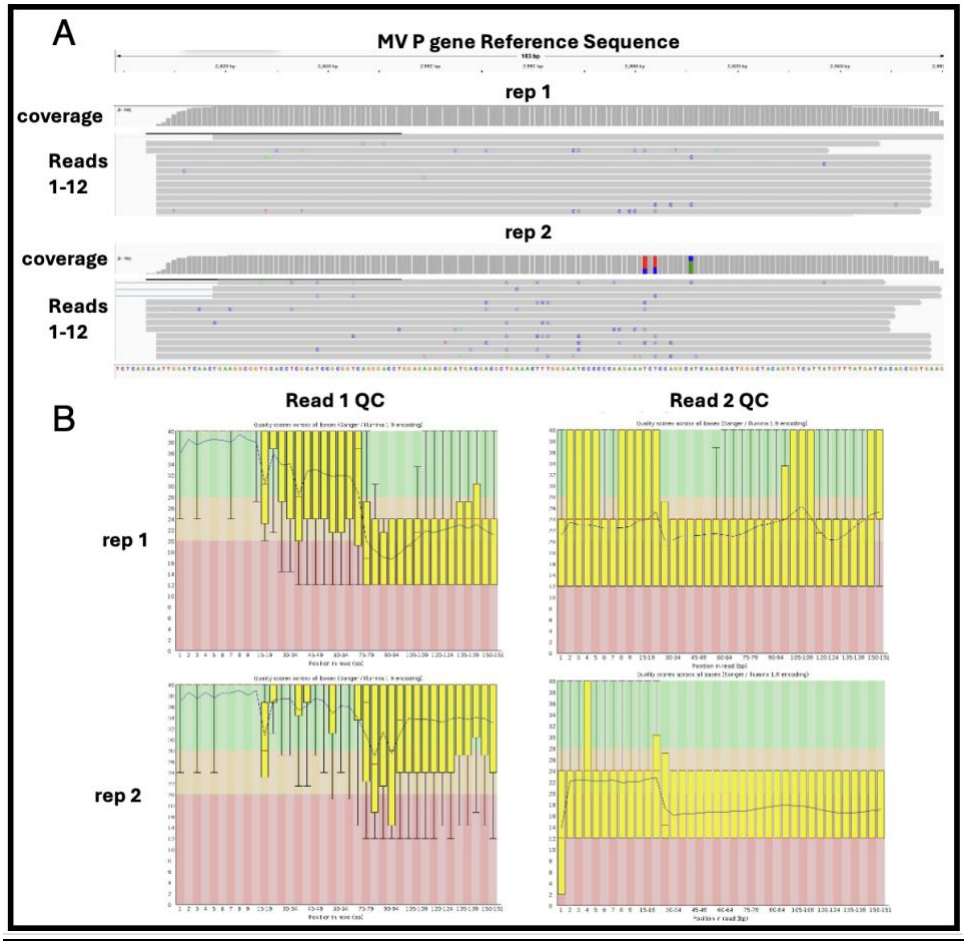


Figure 9 Single-cell mRNA Sequencing of MV P Gene Attenuation Site. (A) P gene mRNA derived from single-cells were assembled by alignment to the MV reference genome. Assembled sequences mapped to a 163 nt region of the P gene, which is known to contain the attenuation site. Mapped sequences showed high genome coverage, though very few reads mapped to the

reference genome. (B) Read quality scores (0 to 40) were higher for read 1 (>20) that contains the barcode sequence, and lower for read 2 (<24) that contains the P gene mRNA sequence.

Discussion

The data presented in this work is preliminary and provides a framework for targeted single-cell transcriptomics of the MV genome. We sequenced a 140 nt region of P gene mRNA from attenuated MV with enhanced function for oncolytic virotherapy [33]-[35],[39]. The goal of this work was to establish a scRNA-seq platform for profiling P gene expression across single MV infected cells. We do so by constructing custom BHBs carrying a GSP sequence that is complementary to MV P gene mRNA (Fig 6). Custom BHBs were co-encapsulated with MV infected Vero cells using droplet microfluidics (Fig 7). A reverse transcription reaction within each drop, containing a single-cell paired with a single BHB, produced P gene cDNA that is barcoded by cell of origin (Fig 7F). This was proceeded with library preparation for Illumina sequencing (Fig 8). We assessed the length and concentration of these libraries for quality control prior to sequencing (Fig 8B and 5.4C). Libraries from two independent replicates were sequenced and produced reads aligning to the expected P gene mRNA region of interest (Fig 9A). Overall, this method expands upon previous scRNA-seq studies that used custom GSPs to target specific RNA regions of interest [22],[23] instead of total mRNA content [18],[19] and provides a new method for examining heterogenous MV infection outcomes during oncolytic virotherapy treatment.

Developing a scRNA-seq method requires optimization of a large parameter space. The first challenge was synthesis of high-quality poly-acrylamide hydrogel beads carrying barcoding oligonucleotide primers. During protocol development, control BHBs were made to validate

procedures for BHB synthesis, sample barcoding, and sequencing library preparation. Control BHBs were initially synthesized with the fluorescent dye FITC-dextran (SI Fig 6A and 5.2B) to increase visibility during BHB re-injection into microfluidic devices. This feature allowed for quantification of BHB loading at different flow rate fractions (Fig 7C, Fig 13). Control BHBs were also made without the incorporation of acrydite-modified DNA, to reduce time and cost of synthesis. This protocol repeatedly produced monodisperse BHBs (Fig 12). However, once combining FITC-dextran and acrydite-modified DNA together in the BHB aqueous phase, we observed inconsistent crosslinking chemistry. Therefore, FITC-dextran was removed from the remaining synthesis runs. Variability in crosslinking chemistry may be due to interactions between the dextran and acrydite that reduced polyacrylamide crosslinking efficacy. Further parameter testing could determine if BHBs can be synthesized with FITC-dextran under different temperature or reagent concentration conditions. This would be advantageous as increased visibility of BHBs allows for easier device tuning during experimentation.

A second set of control BHBs were synthesized to carry a single barcode sequence across all beads. This validated procedures for primer ligation, ensuring that all sequences were present on the beads and functional for their various roles in cDNA synthesis and library preparation. Validation with single-barcode BHBs also reduced the time and cost of assay development. Initially, single-barcode BHBs were ligated with two different GSP sequences, for the MV P gene (SI Fig 6A) and MV L gene (SI Fig 6B). The MV L gene encodes the Large (L) protein that, together with the P protein, forms the MV RNA polymerase complex required for genome replication.

Single barcode BHBs were first used to validate cDNA synthesis and barcoding protocols in bulk. BHBs were split into individual wells of a PCR plate and mixed with all components of the droplet reverse transcription assay. Total well volumes matched that of a pooled droplet sample. Following reverse transcription, BC-cDNA was pooled together and purified. BC-cDNA was detected with custom qPCR primers (Table 4) (SI Fig 6A) and nt length was measured with capillary electrophoresis. This procedure was used to test BHBs against both synthetic RNA templates (Table 3) (Fig 16A) and MV infected cells in bulk (Fig 17A) and drops (Fig 18A). All BC-cDNA products measured between 200-300 nt, in accordance with the expected length.

The library preparation procedure was also first validated with BC-cDNA from previously described single-barcode BHB control experiments (Fig 16-18). Libraries constructed from single-barcode BHBs paired with synthetic RNA templates in bulk measured 489 nt (P gene) and 470 nt (L gene) on average (Fig 16B). These libraries were sequenced with sanger sequencing and contained known barcode and mRNA sequence inserts (Fig 19A). Sequence alignment of both libraries showed shared barcode and Illumina sequences on the 5' and 3' ends, and mismatched mRNA inserts (P or L gene) in the middle of the sequence (Fig 19B). Verification of the barcoding and sequencing protocol with synthetic RNA and reactions in bulk ensured that with high quality BHBs and high concentration RNA templates, the protocol outlined should work for our system. Libraries constructed from AMV infected cells in bulk showed a wider spread of product length (Fig 17B), likely due to the inclusion of intracellular RNA and additional MV RNA in the sample. These samples contained high concentration libraries with an average length of 454 nt (P gene) and 512 nt (L gene)(Fig 20A). Libraries constructed from single-barcode BHBs paired with AMV infected cells in *drops* had a

significantly reduced yield (Fig 18B), with low concentration libraries averaging 538 nt in length (Fig 20B). The decreased yield from bulk to drop control experiments shows the continued need for optimization of the droplet cDNA synthesis and barcoding assay. Unfortunately, none of these libraries could be sequenced on the Illumina platform, as the low sequence diversity of the 5' barcode region prevented the sequencing machine from proper calibration at the start of the run.

Following protocol optimization with single-barcode BHBs, we conducted two independent scRNA-seq experiments to capture P gene mRNA from single infected Vero cells with multi-barcode BHBs. Each BHB was designed to carry a unique 80 nt barcoding sequence. This was accomplished through split-and-pool ligation of four individual 20 nt barcode sequences. Barcodes were prepared in four 96 well plates, allowing for a total of 96^4 (8.49×10^7) unique combinations. In addition to the unique barcodes, each BHB carried sequences for a BclI restriction enzyme site, T7 promoter site, Illumina read 1 site, read specific UMI, and P gene mRNA capture site. Across three independent replicates we observed an average BOP length of 257 nt (Fig 6B). The three replicate BOPs were also verified to contain T7 promoter sequence (located on the 5' end), P gene GSP sequence (located on the 3' end), and therefore all intermediate sequences, by hybridization with complementary Cy5 fluorescent probes (Fig 6C).

Functional BHBs were co-encapsulated with single MV infected Vero cells and a cell lysis / enzyme mixture using a previously designed microfluidic device [22] (Fig 7B). BHB loading did not follow a Poisson distribution (Fig 7C) because tight packing of BHBs is required for single particle loading during re-injection [40]. Cell loading (Fig 7D) did follow a Poisson distribution of 0.5 cells per drop, for a cell suspension of 1×10^6 cells/mL loaded into 60 μm

droplets (Fig 14A and B). Following in droplet reverse transcription for cDNA synthesis and barcoding of P gene mRNA, purified BC-cDNA measured between 200 nt and 300 nt in length (Fig 15B), similar to control experiments. BC-cDNA was also detected by qPCR cycle quantification (C_q) number, amplifying at C_q = 24 and C_q=30 for reps 1 and 2, respectively (Fig 7F). These are lower C_q values than observed during control experiments with single-barcode BHBs (SI Figs 5.7A-5.9A). This may the result of a lower frequency of complete BOPs on multi-barcode BHBs compared to homogenous single-barcode BHBs. Additionally, BC-cDNA yield may be limited by a lower frequency of BHBs paired with infected cells in drops, compared to bulk controls. There also may be limitations due to length of drop-making time, and enzyme exposure to room temperature in the microfluidic tubing, that could have impacted reaction kinetics of the assay in droplets. Increasing droplet collection volume and reducing droplet formation time may increase BC-cDNA yields.

BC-cDNA from two independent replicates was processed for Illumina sequencing using a non-standard library preparation procedure [22]. Unlike typical library preparation procedures, which begin with a cDNA fragmentation step, our procedure requires that cDNA remains intact to preserve attached barcode sequence. Therefore, we must use a method other than fragmentation to standardize library size. This protocol utilizes a second reverse transcription reaction, with a second GSP (Fig 8A iii), to ensure that only a 140 nt region of P gene mRNA is attached to sequencing barcodes. This required the scRNA-seq product to be transferred between multiple nucleic acid types, and multiple rounds of purification between steps, which may have reduced overall product yield. The two sequencing libraries prepared in this work had an average length of 586 nt (rep 1) and 592 nt (rep 2) (Fig 8B), detected at a low concentration of 3.64 ng/µl

(rep 1) and 0.47 ng/µl (rep 2) (Fig 8C). Future directions will include optimization of sequencing library preparation procedures to increase yield. One important addition to the protocol that could help optimize the assay is measuring library nucleotide concentration at intermediate library steps. This may help to adjust reagent concentrations at each step during different independent replicates. Additional suggested troubleshooting and optimization includes suspending nucleotides in TE buffer rather than nuclease-free water, reduction of PCR amplification cycles to limit primer dimers, using a different type of magnetic bead for size selection and purification, running a systematic dilution series with synthetic RNA in drops to identify the limit of detection, decreasing drop-making temperature, and increasing drop-making speed.

P gene sequence assembly showed a low number of high-quality reads mapping to the reference genome, 106 reads for rep 1 and 765 reads for rep 2 (Fig 9A). The aligned reads showed high coverage across a 163 nt region of MV P gene containing our 140 nt region of interest. However, mapped reads were low and thus diversity analysis was not performed. The low percentage of mapped reads is due to low quality read 2 sequences. Read quality scores (0 to 40) were higher for read 1 (>20) that contains the sample barcode sequence, and lower for read 2 (<24) that contains the P gene mRNA sequence (Figure 9B). This is likely a result of low input library concentration (single nanograms per microliter), paired with a high proportion of the library containing primer dimers. Data acquisition is limited by high cost of Illumina sequencing required to troubleshoot the assay. Future optimization of the scRNA-seq protocol should focus on increasing library yield and improving sequencing read quality for population diversity analysis.

This work expands the capabilities of targeted scRNA-seq to MV mRNA regions of interest. Once optimized for *in vitro* Vero cell infection, the MV P gene scRNA-seq platform will be applied to MV *in vivo* infections of PDX models of breast cancer tumor cells. We propose future investigation of the mutational profile of P gene attenuation sites in breast cancer tissue that is susceptible to MV infection compared to breast cancer tissue that was resistant to MV infection. Eventually, our scRNA-seq method will enable targeted scRNA-seq of MV gene expression in tumor cells obtained directly from patients undergoing clinical trials with virotherapy treatment.

Conclusion

scRNA-seq of attenuated MV using droplet microfluidics represents a novel approach for capturing gene expression differences from cell-to-cell during oncolytic virotherapy treatment. Quantification of attenuation site diversity at the single-cell level will increase understanding of subpopulations that contribute to infection suppression phenotypes, and aid in design of better virotherapy treatments that can overcome infection resistant tumor cells. Our method begins with synthesis of custom BHBs for MV P gene mRNA capture. These BHBs, co-encapsulated with single MV infected cells into microfluidic droplets, allow for efficient cDNA synthesis and barcoding of the MV P gene mRNA by cell of origin. This approach promises to offer insights into the diversity of MV gene expression during oncolytic virotherapy treatment and aid in the development of more robust therapeutic strains.

However, this study also faced certain challenges, related to low read quality in the sequencing data, which hindered complete diversity analysis of the MV sequences across single-cells. Despite this, we show preliminary results and control experiments with some successful

sequences for the targeted P gene attenuation site, indicating the potential of this method. Future work will aim to optimize the protocol for improved sequencing yield and enhanced data quality. By refining these techniques, this method could be applied to deeper analyses of MV populations in oncolytic virotherapy and other viral studies, providing a valuable tool for understanding MV gene expression in single-cells and the diversity of MV expression during therapeutic treatment.

References Cited

1. Holland, J. *et al.* Rapid evolution of RNA genomes. *Science* **215**, 1577–1585 (1982).

2. Domingo, E. *RNA Genetics: Volume III: Variability of RNA Genomes*. (CRC Press, London, England, 2020). doi:10.1201/9781351076449.

3. Holland, J. J., De La Torre, J. C. & Steinhauer, D. A. RNA virus populations as quasispecies. *Curr. Top. Microbiol. Immunol.* **176**, 1–20 (1992).

4. Domingo, E. *et al.* Basic concepts in RNA virus evolution. *FASEB J.* **10**, 859–864 (1996).

5. Domingo, E. & Holland, J. J. RNA virus mutations and fitness for survival. *Annu. Rev. Microbiol.* **51**, 151–178 (1997).

6. Domingo, E. Viruses at the edge of adaptation. *Virology* **270**, 251–253 (2000).

7. Vignuzzi, M., Stone, J. K., Arnold, J. J., Cameron, C. E. & Andino, R. Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. *Nature* **439**, 344–348 (2006).

8. Lauring, A. S. & Andino, R. Quasispecies theory and the behavior of RNA viruses. *PLoS Pathog.* **6**, e1001005 (2010).

9. Domingo, E., Sheldon, J. & Perales, C. Viral quasispecies evolution. *Microbiol. Mol. Biol. Rev.* **76**, 159–216 (2012).

10. Andino, R. & Domingo, E. Viral quasispecies. *Virology* **479–480**, 46–51 (2015).

11. Poirier, E. Z. & Vignuzzi, M. Virus population dynamics during infection. *Curr. Opin. Virol.* **23**, 82–87 (2017).

12. Dolan, P. T., Whitfield, Z. J. & Andino, R. Mapping the Evolutionary Potential of RNA Viruses. *Cell Host Microbe* **23**, 435–446 (2018).

13. Dolan, P. T., Whitfield, Z. J. & Andino, R. Mechanisms and concepts in RNA virus population dynamics and evolution. *Annu. Rev. Virol.* **5**, 69–92 (2018).

14. Sanjuán, R. The Social Life of Viruses. *Annu Rev Virol* (2021) doi:10.1146/annurev-virology-091919-071712.

15. Jones, J. E., Le Sage, V. & Lakdawala, S. S. Viral and host heterogeneity and their effects on the viral life cycle. *Nat. Rev. Microbiol.* **19**, 272–282 (2021).

16. Ciuffi, A., Rato, S. & Telenti, A. Single-Cell Genomics for Virology. *Viruses* **8**, (2016).

17. Rato, S., Golumbeanu, M., Telenti, A. & Ciuffi, A. Exploring viral infection using single-cell sequencing. *Virus Res.* **239**, 55–68 (2017).

18. Klein, A. M. *et al.* Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* **161**, 1187–1201 (2015).

19. Macosko, E. Z. *et al.* Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015).

20. Zhou, W.-M. *et al.* Microfluidics applications for high-throughput single cell sequencing. *J. Nanobiotechnology* **19**, 312 (2021).

21. Zilionis, R. *et al.* Single-cell barcoding and sequencing using droplet microfluidics. *Nat. Protoc.* **12**, 44–73 (2017).

22. Chen, K.-Y. *et al.* High-throughput droplet-based analysis of influenza A virus genetic reassortment by single-virus RNA sequencing. *Proc. Natl. Acad. Sci. U. S. A.* **120**, e2211098120 (2023).

23. Saikia, M. *et al.* Simultaneous multiplexed amplicon sequencing and transcriptome profiling in single cells. *Nat. Methods* **16**, 59–62 (2019).

24. Acheson, N. Fundamentals of molecular virology 2nd edition - NLM Catalog - NCBI. https://www.ncbi.nlm.nih.gov/nlmcatalog/101551557 (2011).

25. Anderson, B. D., Nakamura, T., Russell, S. J. & Peng, K.-W. High CD46 receptor density determines preferential killing of tumor cells by oncolytic measles virus. *Cancer Res.* **64**, 4919–4926 (2004).

26. Ong, H. T. *et al.* Oncolytic measles virus targets high CD46 expression on multiple myeloma cells. *Exp. Hematol.* **34**, 713–720 (2006).

27. Condack, C., Grivel, J.-C., Devaux, P., Margolis, L. & Cattaneo, R. Measles virus vaccine attenuation: suboptimal infection of lymphatic tissue and tropism alteration. *J. Infect. Dis.* **196**, 541–549 (2007).

28. Russell, S. J. & Peng, K. W. Measles virus for cancer therapy. *Curr. Top. Microbiol. Immunol.* **330**, 213–241 (2009).

29. Aref, S., Bailey, K. & Fielding, A. Measles to the Rescue: A Review of Oncolytic Measles Virus. *Viruses* **8**, 294 (2016).

30. Msaouel, P. *et al.* Clinical trials with oncolytic measles virus: Current status and future prospects. *Curr. Cancer Drug Targets* **18**, 177–187 (2018).

31. Mühlebach, M. D. Measles virus in cancer therapy. *Curr. Opin. Virol.* **41**, 85–97 (2020).

32. Engeland, C. E. & Ungerechts, G. Measles virus as an oncolytic immunotherapy. *Cancers (Basel)* **13**, 544 (2021).

33. Ohno, S., Ono, N., Takeda, M., Takeuchi, K. & Yanagi, Y. Dissection of measles virus V protein in relation to its ability to block alpha/beta interferon signal transduction. *J. Gen. Virol.* **85**, 2991–2999 (2004).

34. Devaux, P., von Messling, V., Songsungthong, W., Springfeld, C. & Cattaneo, R. Tyrosine 110 in the measles virus phosphoprotein is required to block STAT1 phosphorylation. *Virology* **360**, 72–83 (2007).

35. Devaux, P., Priniski, L. & Cattaneo, R. The measles virus phosphoprotein interacts with the linker domain of STAT1. *Virology* **444**, 250–256 (2013).

36. Russell, S. J. *et al.* Remission of disseminated cancer after systemic oncolytic virotherapy. *Mayo Clin. Proc.* **89**, 926–933 (2014).

37. Galanis, E. *et al.* Oncolytic measles virus expressing the sodium iodide symporter to treat drug-resistant ovarian cancer. *Cancer Res.* **75**, 22–30 (2015).

38. Kemler, I., Ennis, M. K., Neuhauser, C. M. & Dingli, D. In vivo imaging of oncolytic measles virus propagation with single-cell resolution. *Mol. Ther. Oncolytics* **12**, 68–78 (2019).

39. Kurokawa, C. *et al.* Constitutive Interferon Pathway Activation in Tumors as an Efficacy Determinant Following Oncolytic Virotherapy. *J. Natl. Cancer Inst.* **110**, 1123–1132 (2018).

40. Abate, A. R., Chen, C.-H., Agresti, J. J. & Weitz, D. A. Beating Poisson encapsulation statistics using close-packed ordering. *Lab Chip* **9**, 2628–2631 (2009).

CHAPTER SIX

CONCLUSIONS AND OUTLOOK

Conclusions

The results presented in this dissertation expand the use of droplet microfluidics to the study of two clinically relevant negative-strand RNA viruses, IAV and MV. Specifically, we focus on the adaptation of molecular biology assays, like RT-qPCR and next generation sequencing, to a single-cell platform, for nucleic acid quantification and sequencing from IAV and MV RNA genomes and transcripts. Ultimately, the goal of this work was to develop high-throughput methods that can be used to study various features of RNA virus diversity at the single-cell level.

Viruses are dependent upon host cells for replication, and therefore heterogeneity in the host cell population has a strong influence on viral infection dynamics and replication outcomes [1]. The work outlined in chapters 3 and 4 together form a model system for culturing IAV infected host cells in microfluidic droplets and quantifying IAV RNA genome replication from these single encapsulated host cells.

The first platform optimizes cell and virus culture in drops for IAV infection models. In chapter 3, we first examine IAV replication and infectivity across three commonly used cell lines, A549, MDCK and siat7e, using RT-qPCR and plaque assay. This provided baseline measurements for IAV infection in bulk, which were compared to IAV replication and infectivity in our single-cell infection model. To develop the single-cell infection model, we first examined the viability of these three cell lines in 100 μm droplets, after 24 hours of incubation, using live-

dead staining and FACS. Additionally, we measure the stability of droplet interfaces during these incubations, using image analysis. Here, we demonstrate that 100-μm-diameter drops yielded high cell viability (>75%) and drop stability (<3% radii change) following 24 hours of incubation with A549, MDCK, and siat7e cells. However, only A549 cells and MDCK cells showed high levels virus replication in bulk and low variability in single-cell loading in drops.

Our IAV single-cell infection platform was developed for A549 and MDCK cell lines. Measurements of viral RNA via RT-qPCR and infectious progeny via plaque assay were made at 0 and 24 hours from supernatant collection from broken drops and compared to standard bulk infections on well plates. The amount of virus produced from A549 and MDCK cells in bulk was not significantly different than the amount of virus produced from cells in drops. This validated the use of our single-cell infection system for IAV propagation from A549 and MDCK cell lines.

The second platform, described in chapter 4, quantifies the number of extracellular IAV progeny virions that are released from each single-cell encapsulated into microfluidic droplets. This feature of the virus infection cycle is described as viral burst size, and is an important parameter for measuring the success of an infection. The ability to quantify infection success parameters, such as burst size, is an important capability of a model system. By measuring infection success under optimal conditions, we provide a baseline that can be compared against when investigating the impact of variables like host immunity or antiviral drugs that may disrupt IAV replication. To measure IAV burst size from single infected cells, we infected A549 cells with either the H1N1 or H3N2 strain in bulk. These bulk infected cells were encapsulated into 100 μm droplets and incubated for 24 hours to allow for virus replication. After incubation, 1/8th of each droplet volume was split off to sample extracellular IAV progeny virus separate from the

host cell. The fluorescence intensity of each individual droplet was mapped back to a theoretical known starting concentration using an amplification curve mapping model developed specifically for this platform.

Our dqPCR platform for measuring IAV burst size showed increased progeny virus production in the H3N2 population compared to H1N1. Using a negative binomial fit, we estimated that the modeled mean burst sizes for each population were 709 viruses/cell for H3N2 and 358 viruses/cell for H1N1. Both burst size distributions were highly diverse, ranging from ~10^1 to ~10^3 M gene cpd for H1N1 (n = 5,585 drops) and an order of magnitude higher, from ~10^1 to ~10^4 M gene cpd, for H3N2 (n = 9,620 drops). The mean of the H3N2 distribution was 1.0×10^3 M gene cpd. The mean of the H1N1 burst size distribution was 4.8×10^2 M gene cpd. The H3N2 distribution was also less uniform ($G_{H3N2} = 0.705$) compared to H1N1 ($G_{H1N1} = 0.586$). In both viruses, a minor population of infected cells produced a large fraction of progeny virus, but for H3N2, 10% of infected cells produced ≈60% of measured progeny virus, while for H1N1, 10% of infected cells produced ≈40% of progeny virus.

The third and final droplet microfluidics platform described in chapter 5 expands into quantification of viral gene diversity from single infected cells. A custom scRNA sequencing method was designed to capture a small region of interest on MV P gene mRNA. The P gene plays a key immunoregulation role during MV vaccine strain infection that enables their use as an oncolytic virotherapy treatment. The aim of this work was to compare P gene mRNA diversity during oncolytic virotherapy across single-cells, to better understand observed heterogeneity in tumor susceptibility to MV treatment. Instead of using microfluidic droplets to culture and count viral particles, here we employed the technology to synthesize polyacrylamide

hydrogel beads that carry DNA primers for next generation sequencing. When co-encapsulated with single infected cells, these barcoding hydrogel beads facilitate cDNA synthesis and barcoding of viral RNA. We specifically design 30 μm BHBs to carry DNA primers with sequences for Illumina library preparation, a barcode that is unique to each BHB, a molecular ID that is unique to each primer, and a GSP for MV P gene mRNA. A microfluidic device was used to co-encapsulate BHBs with MV infected cells in 60 μm drops with a mixture of enzymes and cell lysis reagents required for cDNA synthesis and barcoding. Drops were broken and cDNA pooled for sequencing library preparation. All libraries were sequenced with Illumina NovaSeq 150-bp paired end sequencing. Consensus sequences were assembled by alignment to the MV vaccine strain derivative reference genome.

We showed reproducible synthesis of ~33 μm diameter poly-acrylamide BHBs carrying BOPs that measured an expected ~257 nt in length. We were able to encapsulate single BHBs into droplets at a high rate (86% of drops), along with single-cells (32% of drops) that followed Poisson distribution. Reverse transcription of P gene mRNA in droplets produced detectable levels of BC-cDNA across two scRNA-seq replicates and multiple control experiments. The BC-cDNA products from both scRNA-seq replicates and control experiments were successfully prepared into Illumina sequencing libraries averaging ~538 nt in total length. Both sanger sequencing and next generation sequencing were used to validate production of sequencing libraries. Low library concentration and subsequent read quality in the scRNA-seq experiments limited any informative diversity analysis from being completed. Therefore, this research currently stands as a framework, requiring further optimization for product yield, for targeted transcriptomics of the MV RNA genome.

The incredible diversity of RNA virus populations, combined with the heterogeneity of host cells they infect, makes studying RNA virus infection quite challenging. Averaging outcomes across RNA virus infection in bulk often misses minor variants and phenotypes that may play an important role in overall population dynamics. Therefore, it is increasingly important to study RNA virus infections at the single-cell level. Droplet microfluidics is a powerful tool for single-cell analysis. By physically isolating individual cells within microfluidic droplets, and manipulating those droplets to conduct various biological assays, we can study RNA virus infection one cell at a time. In this work we established three new droplet microfluidic methods for studying clinically relevant negative-strand RNA viruses, IAV and MV, at the single-cell level. This motivation for this work was to create novel methodologies for high-throughput single-cell analysis of negative-strand RNA virus infection dynamics and population diversity.

Outlook

The field of single-cell virology is rapidly growing, as novel technologies continue emerging to address existing challenges within the field. The remainder of this dissertation will be spent discussing the future directions of single-cell virology, as it pertains to negative strand RNA virus infection as well as the overall outlook of the field.

As illustrated by our work and others, most single-cell analysis tools used to study viral infection rely on microfluidic methods [2]-[4] to measure nucleic acid quantity or sequence [5]. Far fewer platforms are designed to examine other biological material such as proteins or small molecules. Single-cell proteomics relies on protein identification in single cells by mass spectrometry (MS). Two significant platforms have emerged [6], ‘single-cell proteomics by mass

spectrometry' (SCoPE-MS) and nanodroplet processing in one pot for trace samples (nanoPOTS). The SCoPE-MS method first enabled the identification of hundreds of proteins from thousands of single mammalian cells using MS [7]. The nanoPOTS method improved sample preparation protocols and employed protein detection by liquid chromatography MS (LC-MS) to reduce the number of input cells required for high resolution single-cell proteomic analysis [8]. Further optimization of these platforms may include the implementation of label-free and multiplexed methods [6]. Unfortunately, the field of single-cell proteomics has not been widely applied to the study of viral infection, highlighting a need for future research [5]. This application would enable improved understanding of the heterogeneity of both cellular and viral protein modifications during infection.

Future directions of single-cell analysis also include development of spatial analysis platforms. Spatial analysis profiles individual single-cells within original tissue structures without dissociation [9]. This is primarily done using single-cell spatial transcriptomics, which describes gene expression profiles of individual cells according to their location in a tissue [10]. Very few studies have utilized spatial analysis for characterization of viral infection [5]. One study looked at SIV infection in Rhesus Monkeys [11]. Another study examined HSV-1 infection in human brain organoids [12]. A third study profiled spatial transcriptomics in autopsy tissue from human subjects infected with SARS-CoV-2 virus [13]. The lack of single-cell spatial analysis to describe viral infection represents another clear gap in the field.

Existing single-cell virology platforms are limited in their ability to simultaneously measure outcomes from both host cell and virus during infection. This is especially evident in our work quantifying IAV burst size from single-cell infections in microfluidic droplets [14]. To

count the number of IAV particles produced by each infected cell, we must split the virus population from the host cell to measure solely viral RNA. We even use the exclusion of droplets containing cellular RNA from lysate as a quality control measure to ensure we only capture viral RNA. However, to truly understand the context of single-cell data we need to develop ways to pair single-cell measurements from both the host and the virus, before and after infection. New single cell methods for multiplexed analysis, which could be adapted for paired host and virus measurements, continue to emerge [2],[5]. Platforms such as CITE-seq [15] and REAP-seq [16], which combine scRNA-seq with targeted protein quantification, along with PERTURB-seq [17] and ECCITE-seq [18], which combine CRISPR-Cas9 screening with scRNA-seq, are exciting examples of multiplexed/multiomics analysis.

The ideal single cell analysis system for the field of virology would integrate multi-omics measurements (of gene, transcript, and protein diversity) along with paired analysis of host cells and both their incoming and outgoing virus populations. Such a platform would offer a complete snapshot of host-virus infection dynamics at high resolution and throughput, enabling in-depth evolutionary studies of how virus and host populations change over time in response to selective pressures. This system would be invaluable for studying differential responses to drug and antiviral treatments, and allow for rapid analysis and refinement of the most effective candidates. Continued development of tools like these are essential for accelerating research aimed at treating and preventing the spread of dangerous infectious diseases in humans [4]. Especially as global climate change begins to shift the distribution of novel pathogenic organisms, bringing them closer to human populations, it is paramount that we continue to invest in research tools that can address these emerging challenges.

References Cited

1. Jones, J. E., Le Sage, V. & Lakdawala, S. S. Viral and host heterogeneity and their effects on the viral life cycle. *Nat. Rev. Microbiol.* **19**, 272–282 (2021).

2. Matuła, K., Rivello, F. & Huck, W. T. S. Single-Cell Analysis Using Droplet Microfluidics. *Adv Biosyst* **4**, e1900188 (2020).

3. Zhou, W.-M. *et al.* Microfluidics applications for high-throughput single cell sequencing. *J. Nanobiotechnology* **19**, 312 (2021).

4. Lehnert, T. & Gijs, M. A. M. Microfluidic systems for infectious disease diagnostics. *Lab Chip* **24**, 1441–1493 (2024).

5. Bost, P. & Drayman, N. Dissecting viral infections, one cell at a time, by single-cell technologies. *Microbes Infect.* **26**, 105268 (2024).

6. Truong, T. & Kelly, R. T. What's new in single-cell proteomics. *Curr. Opin. Biotechnol.* **86**, 103077 (2024).

7. Budnik, B., Levy, E., Harmange, G. & Slavov, N. SCoPE-MS: mass spectrometry of single mammalian cells quantifies proteome heterogeneity during cell differentiation. *Genome Biol.* **19**, 161 (2018).

8. Zhu, Y. *et al.* Nanodroplet processing platform for deep and quantitative proteome profiling of 10-100 mammalian cells. *Nat. Commun.* **9**, 882 (2018).

9. Moffitt, J. R., Lundberg, E. & Heyn, H. The emerging landscape of spatial profiling technologies. *Nat. Rev. Genet.* **23**, 741–759 (2022).

10. Ståhl, P. L. *et al.* Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* **353**, 78–82 (2016).

11. Jiang, S. *et al.* Combined protein and nucleic acid imaging reveals virus-dependent B cell and macrophage immunosuppression of tissue microenvironments. *Immunity* **55**, 1118-1134.e8 (2022).

12. Rybak-Wolf, A. *et al.* Neurodegeneration in human brain organoids infected with herpes simplex virus type 1. *bioRxiv* (2021) doi:10.1101/2021.03.05.434122.

13. Butler, D. *et al.* Shotgun transcriptome, spatial omics, and isothermal profiling of SARS-CoV-2 infection reveals unique host responses, viral diversification, and drug interactions. *Nat. Commun.* **12**, 1660 (2021).

14. Thomas, M. M. *et al.* Influenza A viral burst size from thousands of infected single cells using droplet quantitative PCR (dqPCR). *PLoS Pathog.* **20**, e1012257 (2024).

15. Stoeckius, M. *et al.* Simultaneous epitope and transcriptome measurement in single cells. *Nat. Methods* **14**, 865–868 (2017).

16. Peterson, V. M. *et al.* Multiplexed quantification of proteins and transcripts in single cells. *Nat. Biotechnol.* **35**, 936–939 (2017).

17. Dixit, A. *et al.* Perturb-seq: Dissecting molecular circuits with scalable single-cell RNA profiling of pooled genetic screens. *Cell* **167**, 1853-1866.e17 (2016).

18. Mimitou, E. P. *et al.* Multiplexed detection of proteins, transcriptomes, clonotypes and CRISPR perturbations in single cells. *Nat. Methods* **16**, 409–412 (2019).

APPENDICES

APPENDIX A

SUPPLEMENTAL INFORMATION FOR CHAPTER 4

Supplementary Information

Influenza A Viral Burst Size from Thousands of Infected Single Cells Using Droplet Quantitative PCR (dqPCR)

Geoffrey K. Zath[1,2¶], Mallory M. Thomas[1,3,4¶], Emma Kate Loveday[1,2], Dimitri A. Bikos[1,2], Steven Sanche[5], Ruian Ke[5], Christopher B. Brooke[6], Connie B. Chang[3*]

[1]Center for Biofilm Engineering, Montana State University, Bozeman, Montana, United States of America
[2]Department of Chemical and Biological Engineering, Montana State University, Bozeman, Montana, United States of America
[3]Department of Physiology and Biomedical Engineering, Mayo Clinic, Rochester, Minnesota, United States of America
[4]Department of Microbiology and Immunology, Montana State University, Bozeman, Montana, United States of America
[5]Theoretical Biology and Biophysics, Los Alamos National Laboratory, Los Alamos, New Mexico, United States of America
[6]Department of Microbiology, University of Illinois, Urbana-Champaign, Urbana, Illinois, United States of America

[¶]Authors contributed equally.

*Corresponding author

Email: chang.connie@mayo.edu

Table of Contents

Supplementary materials and methods

(M1) RT-qPCR Template and Primer Sequences. Detection of IAV and host cells was facilitated by reverse transcription quantitative polymerase chain reaction (RT-qPCR) of IAV matrix protein genomic RNA (M gene) and cellular β-actin mRNA. M gene RNA was amplified with forward/reverse primers and a Taqman probe (FAM/BHQ1) that target a highly conserved region shared across IAV strains (1) (Table S1). β-actin mRNA was amplified with forward/reverse primers and a Taqman probe (Cy5/BHQ2) shown to be robust against multiple eukaryotic cell lines (2) (Table S1). All primers and probes were purchased from Eurofins Operon as 100 μM stocks. Known concentrations of the template sequences were used as positive controls for RT-qPCR validation throughout this work. The M gene control was first constructed as linear dsDNA (IDT) containing a T7 promoter (underlined), forward and reverse primer binding sites (italicized), and a probe binding site (bold) (Table S1). M gene RNA was *in vitro* transcribed (IVT) from the ssDNA gblock using the MEGAscript T7 RNA Synthesis Kit (Ambion, #AM1333), and purified over a GE Illustra Sephadex G-50 NICK column. The β-actin control (Table S1) was constructed as a dsDNA plasmid (pCAG-mGFP-Actin, Addgene #21948), also containing forward and reverse primer binding sites (italicized) and a probe binding site (bold), which was directly amplified by PCR. The concentrations (copies/μL) of template sequence controls in working stocks were quantified with a NanoDrop spectrophotometer. For droplet experiments, template sequence concentrations in copies per drop (cpd) were calculated by multiplying copies/μL by the volume (μL) of a 50 μm (6.54×10^{-5} μL) or 100 μm (or 5.24×10^{-4} μL) diameter drop.

(M2) Microfluidic Device Fabrication. Microfluidic devices were fabricated in polydimethylsiloxane (PDMS) (Sylgard 184) using soft lithography. Three device geometries were designed using AutoCAD (Autodesk): a flow-focusing device for encapsulation of infected cells in 100 μm drops, a split-and-merge device used to isolate replicated virus from the host cell and merge with RT-qPCR mix (3), and a flow-based imaging device (4) for drop fluorescence detection. Device master molds were fabricated by patterning SU-8 photoresist (Microchem, SU-8 3050) on silicon wafers (University Wafer, #447) with photolithography. PDMS was prepared at a

10:1 mass ratio of polymer to cross-linking agent and poured onto the device master molds. Air was purged from the uncured PDMS by placing the filled mold in a vacuum chamber for at least 1 hr. The PDMS was cured in an oven at 55 °C for 24 hrs and then removed from the mold with a scalpel. Inlet and outlet ports were punched into PDMS slabs with a 0.75 mm diameter biopsy punch (EMS Rapid-Core, Electron Microscopy Sciences). The split-and-merge device was comprised of two layers bonded together after plasma treatment (Harrick Plasma, PDC-001) for 30 s at medium power and 700 mTorr oxygen pressure. The double-layer PDMS split-and-merge device and single-layer PDMS flow-focusing device were each bonded to 2 × 3 in glass slides (VWR micro slides, #48382-179) after plasma treatment for 60 s at high power and 400 mTorr oxygen pressure. Microelectrodes were embedded in the split-and-merge device by injecting molten solder (Indium Corporation of America, #53307) into the electrode channels, and terminated with a pin terminal (Phoenix Contact, #1945151). Both devices were made hydrophobic by flowing Aquapel (Pittsburgh Glass Works) through the channels, followed by blowing the channels with air passed through a 0.2-µm filter (GVS ABLUO 25 mm, Fisher Scientific) before baking the devices in an oven at 55 °C for 1 hr for further drying.

(M3) Bulk RT-qPCR. Bulk RT-qPCR reactions were completed using the SuperScript III Platinum One-Step RT-qPCR kit (Invitrogen 11732020) with a final reaction volume of 25 µL. Working stocks of M gene primers and FAM TaqMan probe (Table S1) were prepared at 25 µM and 10 µM, respectively. The bulk RT-qPCR master mix contained final concentrations of 400 nM primers, 200 nM probe, 0.05 µM ROX reference dye, 2.0 mM $MgSO_4$, 1.0% w/v Tween-20 (Calbiochem 655204-100mL), 0.8 µg/µL BSA (Fisher BP675-1), 1.0 M betaine (Sigma B0300-1VL), 0.32 U/µL SUPERase RNase Inhibitor (Invitrogen AM2694), and 1 µL of SuperScript III RT/Platinum Taq mix (Invitrogen 11732020). The additives (Tween-20, BSA, betaine) increase stability of dqPCR reactions, as determined by our previously published protocol (5). Thermocycling was performed using a standard qPCR machine (QuantStudio 7, Applied Biosystems) with the following conditions: 1 cycle for 30 mins at 60 °C, 1 cycle for 2 mins at 95

°C, and 40 cycles between 15 s at 95 °C and 1 min at 60 °C. Reference amplification curves for the bulk RT-qPCR assay were constructed from a dilution series of M gene IVT RNA (Table S2).

(M4) Droplet Quantitative PCR (dqPCR) of Isolated Progeny Viruses. A multiplexed dqPCR assay was used to amplify both IAV M gene RNA and cellular β-actin mRNA in drops containing progeny virus. The dqPCR master mix contained the same reagent concentrations as the bulk RT-qPCR, except for an increased concentration of ROX reference dye at 0.15 μM. Working stocks of the M gene primers and FAM TaqMan probe, as well as the β-actin primers and Cy5 TaqMan probe (Table S1), were prepared at concentrations of 25 μM and 10 μM, respectively. Droplets containing infected cells were re-injected into a split-and-merge device (3) at a flow rate of 500 μL/hr with oil (3.0 wt% RAN surfactant in HFE 7500) flowing at 2000 μL/hr to space them apart. In the device, a split junction was used to isolate progeny viruses from the host cell. Approximately 1/8th volume of the drop containing progeny virus was split and sent to the merge junction. To separate split drops before the merge junction, a second spacer oil (3.0 wt% RAN surfactant in HFE 7500) was injected into the device at a flow rate of 500 μL/hr. At the merge junction, the RT-qPCR master mix was injected into the split drop at a flow rate of 438 μL/hr. Merging was facilitated by an embedded microelectrode that provided a constant 25 kHz, 200 V square wave signal to destabilize the drop interface and allow merging with the RT-qPCR mix. The microelectrode was connected to a high voltage amplifier (Trek Model, #2220-CE) controlled by a custom LabVIEW program. Merged drops were collected in PCR tubes for 2.5 min each to sample approximately 20 μL of drops. PCR tubes were stored on a cold block until ready for thermocycling. Drops were thermocycled in a standard qPCR machine (QuantStudio 7, Applied Biosystems) with the following conditions: 1 cycle for 30 min at 60 °C, 1 cycle for 2 min at 95 °C, and 40 cycles between 15 sec at 95 °C and 1 min at 60 °C. Drops were removed from the thermocycler at cycle numbers N = 1 and N = 40, to measure baseline and terminal fluorescence, as well as multiple intermediate cycle numbers falling within the exponential to linear region of PCR amplification curves.

Reference amplification curves for the dqPCR assay were constructed by encapsulating 10^6 copies/μL of M gene IVT RNA and β-actin plasmid (Table S1) into 50 μm drops using a flow-focusing device, which corresponds to 10^2 cpd of M gene and β-actin. Drops containing M gene or β-actin template sequence controls were collected in 3 mL syringes for 36 min to generate 600 μL of drops and kept on ice until thermocycling with the above conditions. The method for constructing reference amplification curves from dqPCR of template sequence controls is further described in **SI M8.**

(M5) Drop Fluorescence Imaging. As a positive control, thermocycled drops were imaged to capture five fields of view on an inverted epifluorescence microscope (Nikon Ti2-E) at 10× magnification (NA 0.3). Brightfield and fluorescence images were captured with a sCMOS camera (Hamamatsu, ORCA-Flash 4.0 v3) for each reporter fluorescence, FAM TaqMan for M gene RNA (FITC channel), Cy5 TaqMan probe for β-actin RNA (Cy5 channel), as well as for the ROX reference dye (Texas Red channel). Drop fluorescence intensities are reported as ΔR_N, which is the ratio of the M gene (FAM) or β-actin (Cy5) reporter florescence to the reference dye fluorescence (ROX) at each sampled cycle number (N), normalized to the baseline at N = 1. Here, $R_N = (F_{N,\,FAM}$ or $F_{N,\,Cy5})/F_{N,\,ROX}$ and $\Delta R_N = R_N - R_{N=1}$.

(M6) Drop Flow-based Fluorescence Detection. Thermocycled drops were injected into a continuous-flow microfluidic detection device (4) at a flowrate of 200 μL/hr, along with fluorinated spacer oil (HFE 7500) injected at a flowrate of 800 μL/hr. Drop fluorescence intensity (ΔR_N) was detected in high-throughput using a Nikon Ti-U inverted microscope custom-modified with three lasers, a set of dichroic mirrors, and two photomultiplier tubes (PMTs, Hamamatsu H10723-20). The beams of the 22 mW 488 nm laser (Thorlabs MCLS1), 25 mW 561 nm laser (Cobolt Jive), and 20 mW 642 nm laser (Thorlabs MCLS1) were aligned and coupled into the backport of the microscope where they were focused to a spot by the 40× objective (NA 0.60). The flowing of drops across the laser spot resulted in fluorescence detection by three PMTs split into three channels using dichroic and bandpass filters, Ch1: 520/40 nm (green), Ch2: 600/40 nm (red), and

Ch3: 670/10 nm (far-red). A field programmable gate array (FPGA, National Instruments NI-7852R) was used to control the PMT gains and record fluorescence measurements using LabVIEW 2015. A custom MATLAB (R2020a) script was used to process and analyze the drop fluorescence detection data.

(M7) Cycle Threshold (*Ct*) Method of Creating Standard Curves for RT-qPCR. In standard bulk RT-qPCR, the nucleic acid template concentration of an unknown sample is determined by comparing its cycle threshold (*Ct*) value to a standard curve made from a dilution series of known template concentrations. The standard curve follows Eq. S1:

$$\boldsymbol{Ct = m * log_{10}(C_{RNA}) + b} \quad \textbf{(Eq. S1)}$$

Cycle threshold (*Ct*) is the PCR cycle number at which template amplification fluorescence rises above background fluorescence, C_{RNA} is the known template concentration, m is the slope of the *Ct* versus $\log_{10}(C_{RNA})$ line, and b is the y-intercept of the line.

This method assumes that PCR efficiency at the *Ct* value (E_{Ct}) is constant, regardless of starting template concentration. An optimized PCR assay should have an E_{Ct} between 90 – 110% for accurate quantification of template concentration. E_{Ct} is calculated using m of the *Ct* versus $\log_{10}(C_{RNA})$ line, as shown in Eq. S2:

$$\boldsymbol{E_{Ct} = 10^{\frac{1}{m}} - 1} \quad \textbf{(Eq. S2)}$$

For an ideal assay with E_{Ct} = 100%, the Ct values of a 10-fold dilution series are spaced approximately 3.3 cycles apart ($m \simeq -3.3$).

(M8) Sigmoidal Curve Fitting using PCR Efficiency (SCF-E). To track RNA amplification during dqPCR, we reconstructed continuous RT-qPCR amplification curves from discontinuous drop fluorescence measurements, as shown in Fig. 2A. Amplification curves can be fit to discontinuous drop fluorescence measurements using a sigmoidal curve fitting (SCF) model. Here, we employ SCF that relates reaction efficiency to fluorescence intensity to accurately model the shape of RT-qPCR amplification curves. This method is called "Sigmoidal Curve Fitting Using PCR Efficiency," or SCF-E. We used SCF-E to generate reference amplification curves of a known RNA template concentration during dqPCR.

RT-qPCR amplification curves can have an asymmetric shape due to a decrease in PCR reaction efficiency after each cycle number. At early cycles, PCR efficiency is near 100%, relating to an exact doubling of nucleic acid template sequence at each cycle, and approaches 0% at later cycles as PCR reagents become consumed (6). This relationship leads to a steeper slope at the start of the curve during exponential amplification and a flatter slope near the final plateau phase (7). To model asymmetric amplification curves from fluorescence data alone, SCF methods typically set a limit on the number of cycles that are fit (8) or add asymmetric fitting terms (9,10). Alternative methods first model PCR efficiency with a sigmoidal (11) or bilinear (7) equation, and use this to construct a fluorescence amplification curve. In our SCF-E model, asymmetric RT-qPCR amplification curves are fit to experimental fluorescence measurements using a four-parametric sigmoid function that includes a PCR efficiency parameter, E_N, as shown in Eq. S3:

$$E_N = \frac{E_{max}}{1 + e^{\left(\frac{N - N_{0.5}}{k}\right)}} + E_{min} \qquad \textbf{(Eq. S3)}$$

E_N is the PCR efficiency at cycle N, E_{max} is the maximum PCR efficiency in the range of 0.9 to 1, E_{min} is the minimum PCR efficiency in the range of 0 to 0.1, and $N_{0.5}$ is the cycle at which E_N equals 0.5. The parameter k determines the shape of the amplification curve, with larger values

flattening the curve and smaller values steepening the curve. An example of the PCR efficiency curve is illustrated in Fig. S2 (blue dashed curve).

During RT-qPCR, RNA amplification is detected with a complementary fluorescent probe during thermocycling. The rate of fluorescence accumulation during RT-qPCR is determined by the reaction efficiency, as described by Eq. S4:

$$F_{N+1} = F_N(1 + E_N) \quad \textbf{(Eq. S4)}$$

F_N is the fluorescence intensity of the amplification curve at cycle N, and F_{N+1} is the fluorescence intensity of the amplification curve at cycle $N+1$. The fluorescence amplification curve is also illustrated in Fig. S2 (blue solid curve).

Eqs. S3 and S4 are combined to yield Eq. S5, the **SCF-E model**:

$$F_{N+1} = F_N\left(1 + \frac{E_{max}}{1 + e^{\left(\frac{N - N_{0.5}}{k}\right)}} + E_{min}\right) \quad \textbf{(Eq. S5)}$$

To generate an RT-qPCR amplification curve of F_N at all PCR cycle numbers (N = 1 to 40), we vary the efficiency parameters in Eq. S3 (E_{max}, E_{min}, $N_{0.5}$, and k) to find the best fit to our experimental F_N values, measured at only a few PCR cycle numbers. Thirty values of each of the four efficiency parameters are used to test 8.1×10^5 (30^4) curve fits. The set of parameters that yield the best fit (highest R^2 value) to experimental F_N measurements are used to construct the RT-qPCR amplification curve. An initial input for F_N of any non-zero integer is chosen for the fit; here, we used F_N = 1. A calibration factor (CF), shown in Eq. S6, is used to normalize F_N at N =

40 in the SCF-E model curve ($F_{endpoint,\ model}$) to the experimentally measured F_N at N = 40 ($F_{endpoint,\ measured}$):

$$CF = F_{endpoint,measured}/F_{endpoint,model} \quad \textbf{(Eq. S6)}$$

(M9) Amplification Curve Library (ACL) Method of Constructing Standard Curves for RT-qPCR. We generated a high-resolution dilution series of 1000 RNA template concentrations to construct standard curves for conversion of drop fluorescence to RNA concentration. This is performed by first generating a single reference amplification curve for a known IAV M gene RNA concentration (10^2 cpd) (Fig. 2A), that is fit using the SCF-E model (Eq. S5). The reference amplification curve is then translated along its x-axis by varying the SCF-E $N_{0.5}$ parameter 1000 times, between cycles 1 to 40, to build what we call an "Amplification Curve Library" (ACL) of 1000 virtual curves (Fig. 2B, plot with subset of 10 curves). The ACL is then used to construct a virtual standard curve from a single PCR cycle number (Fig. 2C). This method for generating standard curves is described in detail below in three sections.

(1) Translating a Single SCF-E Reference Curve into 1000 Virtual Amplification Curves: The ACL is constructed by expanding a single reference amplification curve, fit with Eq. S5, into 1000 virtual curves. This is done by substituting $N_{0.5,\ ref}$, fit from Eq. S5, with $N_{0.5,\ vir}$, to make Eq. S7:

$$F_{N+1} = F_N\left(1 + \frac{E_{max}}{1 + e^{\left(\frac{N - N_{0.5,vir}}{k}\right)}} + E_{min}\right) \quad \textbf{(Eq. S7)}$$

The translation of the reference curve into virtual curves using Eq. S7, is illustrated in Fig. S3. Here, F_N and E_N are plotted on the left and right y-axes, N on the x-axis (Fig. S3). The solid blue

curve is an example of a reference curve fit using the SCF-E model (Eq. S5) and the dashed blue curve is the corresponding PCR efficiency curve, as previously described (Fig. S2). The reaction efficiency parameters in Eq. S7 (E_{max}, E_{min}, k) are taken from the reference curve fit (Eq. S5) and are constant across all the virtual curves. In the ACL, each virtual curve is defined by a different cycle number at which PCR reaction efficiency equals 0.5 ($N_{0.5,\ virt}$). We input 1000 evenly spaced $N_{0.5,\ virt}$ values, between cycle numbers N = 1 to 40, into Eq. S7 to produce 1000 unique virtual curves. Two virtual curves are displayed in Fig. S3, to illustrate that $N_{0.5,\ virt} < N_{0.5,\ ref}$ (pink curve) is associated with virtual RNA concentrations ($C_{RNA,\ virt}$) greater than the reference ($C_{RNA,\ ref}$). Conversely, if $N_{0.5,\ virt} > N_{0.5,\ ref}$ (yellow curve) then $C_{RNA,\ virt} < C_{RNA,\ ref}$.

(2) Calculating RNA Template Concentration of the Virtual Amplification Curves:

$C_{RNA,\ virt}$ is determined by its x-axis distance from $C_{RNA,\ ref}$. Since the x-axis of the ACL is cycle number, we can relate the virtual and reference curves using the cycle numbers at which PCR efficiency equals 0.5, or $N_{0.5,\ virt}$ and $N_{0.5,\ ref}$. The $N_{0.5,\ virt}$ is a fixed array of 1000 intervals between cycle numbers N = 1 to 40, as previously mentioned (Eq. S7), while $N_{0.5,\ ref}$ is a single cycle number on the reference curve (fit from Eq. S5). The relationship between an amplification curve's $N_{0.5}$ value and its C_{RNA} value yields a line that follows Eq. S8 below, which is analogous to the Ct method (Eq. S1), with $N_{0.5}$ used in place of Ct.

$$N_{0.5} = m * log_{10}(C_{RNA}) + b \qquad \textbf{(Eq. S8)}$$

We expand this equation for $N_{0.5,\ virt}$ and $N_{0.5,\ ref}$ and take the difference with Eq. S9:

$$N_{0.5,virt} - N_{0.5,ref} = [m * log_{10}(C_{RNA,virt}) + b] - [m * log_{10}(C_{RNA,ref}) + b] \qquad \textbf{(Eq. S9)}$$

Eq. S9 is rearranged in Eqs. S10-11 in order to solve for $C_{RNA,\,virt}$ in Eq. 12:

$$log_{10}(C_{RNA,virt}) = log_{10}(C_{RNA,ref}) + \frac{N_{0.5,virt} - N_{0.5,ref}}{m} \quad \textbf{(Eq. S10)}$$

$$f = log_{10}(C_{RNA,ref}) + \frac{N_{0.5,virt} - N_{0.5,ref}}{m} \quad \textbf{(Eq. S11)}$$

$$C_{RNA,virt} = 10^{f} \quad \textbf{(Eq. S12)}$$

Note that the value of $C_{RNA,\,virt}$ varies with m from Eq. S8 when used in Eq. S11. To calculate m, Eq. S2 is rearranged into Eq. S13:

$$m = \frac{1}{log_{10}(E_N + 1)} \quad \textbf{(Eq. S13)}$$

m is dependent upon E_N of the reference curve at a single cycle number, calculated with Eq. S3. As E_N increases, m decreases, and the range of $C_{RNA,\,virt}$ values in the ACL widens (Fig. S4). To determine the E_N constant used to calculate m in Eq. S13, we empirically compared standard curves from libraries built with different E_N values **(SI R3)**. Briefly, E_N corresponding to $N = N_{0.5} - 3$ was used to determine virtual RNA concentration for all ACLs used in this work.

(3) Constructing a Standard Curve from the ACL: To determine the RNA concentrations of unknown samples based on their fluorescence intensity, we utilize the ACL (Fig. S5A) to construct standard curves that relate C_{RNA} to F_N (Fig. S5B). The standard curve is constructed by selecting one PCR cycle number in the ACL (Fig. S5A, vertical line) and relating

305 F_N at that cycle number to $C_{RNA,\ vir}$ of curves intersecting the line. The usable region of the
306 standard curve (Fig. S5B, grey box) generally corresponds to the exponential to linear region of
307 the virtual amplification curves (Fig. S5A, grey box). Outside of these regions, F_N has either not
308 risen above baseline fluorescence at early cycle numbers or has reached saturation at late cycle
309 numbers. Threshold values are set for F_N to define these usable regions, where the upper
310 threshold is set at 1 standard deviation below the F_N at $N = 40$, and the lower threshold is set as
311 the 99th percentile of the F_N at $N = 1$, similar to the thresholds generally set by commercial qPCR
312 machines.
313

314 ## Supplementary results

315 **(R1) Comparing M Gene Abundance During Bulk and Drop Infections.** To demonstrate that
316 encapsulation in drops does not affect IAV infection dynamics (12) during our burst size
317 experiments, we compared the M gene RNA abundance in drop and bulk infections using a bulk
318 RT-qPCR assay (Fig. S1). For bulk infections, we collected both the supernatant and cellular
319 monolayer from each well and prepared them for bulk RT-qPCR. The 0 hpi bulk samples were
320 frozen overnight at -80 °C while the 18 hpi bulk samples were incubated overnight at 37 °C.
321 Preparation of drop infections for bulk RT-qPCR involved freezing a 300 µL sample of drops at -
322 80 °C for 30 mins to break the emulsion. 200 µL of the broken emulsion was collected for further
323 processing. Collected bulk and drop infections were clarified by centrifugation at 500 × *g* for 5
324 min, and the resulting supernatant was sampled for RT-qPCR targeting the IAV M gene. For both
325 strains, there was greater virus production in drops compared to bulk infection (Fig. S1, 18 hpi),
326 when measured with a two-sample Student's t-test ($p << 0.05$). The relative increase in virus
327 production, between drop and bulk infection, was 6.8× for H3N2 and 2.4× H1N1 infections. On
328 average, H3N2 production in drops was 5.0× greater than H1N1 ($p << 0.05$).
329

330 **(R2) Validating SCF-E with Bulk RT-qPCR of IAV M Gene.** Using bulk RT-qPCR amplification
331 of the IAV M gene across five orders of magnitude (Table S2, 10 fold dilutions of 2.62×10^4 to
332 2.62×10^9 copies/µL), we validated the use of our SCF-E model (Fig. S6A) compared to the

standard SCF model (Fig. S6B) (8,13). SCF-E fit parameters are listed in Table S3. The SCF-E model provided a better fit across all six RNA concentrations compared to SCF (Table S3, R^2 of SCF-E vs. SCF). The differences in fit were most pronounced in the exponential and plateau regions of RT-qPCR amplification curves, as demonstrated in Fig. S6C for 2.62×10^7 copies/µL (magnified insets).

(R3) Setting a PCR Efficiency Constant (E_N) for the ACL. To calculate $C_{RNA,\ virt}$ in the ACL, we needed to determine the E_N constant used to calculate m in Eq. S13. E_N is the PCR efficiency at each cycle number (N = 1 to 40) in the SCF-E reference curve (Eq. S5) and m is the slope of the C_{RNA} vs. $N_{0.5}$ line (Eq. S8). To determine which of the 40 possible reference curve E_N values to use in Eq. S13, we constructed ACLs with different E_N values and compared the dynamic range of their standard curves (Table S4). A schematic of the iterative process for determining E_N at different N values is displayed in Fig. S7.

We first tested E_N at $N = N_{0.5} - 0$, using a SCF-E reference curve of 10^7 copies/µL of IAV M gene RNA amplified by bulk RT-qPCR. This yielded E_N = 0.514 and corresponding m = 5.55 (Table S4). These values were used to build an ACL, from which we constructed a standard curve at cycle 20. The standard curve was used to convert F_N to C_{RNA} for a known dilution series of IAV M gene RNA (Table S2, 3 replicates of 21 concentrations between 2.62×10^4 to 2.62×10^9 copies/µL) amplified by bulk RT-qPCR. To determine the accuracy of the standard curve, we performed a linear regression between measured ($C_{RNA,\ ACL}$) and expected ($C_{RNA,\ expected}$) RNA concentrations (Fig. S8, $N = N_{0.5} - 0$, R^2 = 0.886). Samples in the dilution series with a $C_{RNA,\ ACL}$ value within 2-fold change of $C_{RNA,\ expected}$ were included in the dynamic range of the standard curve (Table S4). For E_N at $N = N_{0.5} - 0$, the dynamic range of the cycle 20 standard curve was small, a value of 1.60 $\log_{10}$ RNA copies/µL (Table S4, dynamic range of 3.28×10^6 to 1.31×10^6 copies/µL). Thus, to widen the dynamic range, we tested higher E_N values at five cycles preceding $N_{0.5}$ (Table S4, $N = N_{0.5} - 1$, $N = N_{0.5} - 2$, $N = N_{0.5} - 3$, $N = N_{0.5} - 4$, and $N = N_{0.5} - 5$). The ACL standard curve with the widest dynamic range and highest degree of accuracy (R^2)

361 between $C_{RNA,\ ACL}$ and $C_{RNA,\ expected}$ was built with $N = N_{0.5} - 3$, where $E_N = 0.745$. For this ACL, the
362 dynamic range of the standard curve was 4.70 $\log_{10}$ RNA copies/µL (Table S4, dynamic range of
363 2.62×10^4 to 1.31×10^9 copies/µL) with $R^2 = 0.990$ between $C_{RNA,\ ACL}$ and $C_{RNA,\ expected}$. Therefore,
364 the E_N corresponding to $N = N_{0.5} - 3$ was used to build all amplification curve libraries in this work.
365
366 **(R4) Validating ACL Standard Curves for Multiple PCR Cycle Numbers.** To pool $C_{RNA,\ ACL}$
367 values sampled at multiple PCR cycle numbers, we validated that ACL standard curves
368 converted F_N to $C_{RNA,\ ACL}$ within an acceptable error range of 2-fold change from $C_{RNA,\ expected}$, at
369 any cycle number. This was done using a known dilution series of IAV M gene RNA amplified by
370 bulk RT-qPCR (Table S2, 3 replicates of 21 concentrations between 2.62×10^4 to 2.62×10^9
371 copies/µL). Each ACL was built from a SCF-E reference curve of 2.62×10^7 M gene copies/µL.
372 Standard curves were constructed at ACL $N = 17$ to 23, and used to converted F_N to $C_{RNA,\ ACL}$ for
373 the above dilution series. These standard curve cycle numbers were chosen as they typically fall
374 in the linear region of our M gene reference curves. To determine how well $C_{RNA,\ ACL}$ matched to
375 $C_{RNA,\ expected}$, we performed a linear regression on these values (Fig. S9). Samples in the dilution
376 series whose $C_{RNA,\ ACL}$ fell within a 2-fold change of $C_{RNA,\ expected}$ were included in the dynamic
377 range for each standard curve (Table S5). We found a wide dynamic range (Table S5, dynamic
378 ranges from 4.00 to 4.70) and a strong linear relationship between the $C_{RNA,\ ACL}$ and $C_{RNA,\ expected}$
379 (Fig. S9, R^2 ranging from 0.800 to 0.991) for standard curves from all tested cycle numbers.
380
381 We validate that the performance of ACL standard curves at different cycle numbers is
382 comparable to that of *Ct* standard curves used in typical bulk RT-qPCR. With a *Ct* standard
383 curve, we convert F_N to $C_{RNA,\ Ct}$ for the same dilution series of IAV M gene RNA (Table S2). This
384 conversion resulted in a wide dynamic range of 5 orders of magnitude and a strong linear
385 response between $C_{RNA,\ Ct}$ and $C_{RNA,\ expected}$ of $R^2 = 0.991$ (Fig. S9, under *Ct* plot; Table S5 under
386 *Ct* method).
387

To further quantify the variability between standard curves from different ACL cycle numbers, we calculated the coefficient of variation (*CV*) of $C_{RNA,\ ACL}$ replicates (Fig. S10A) and percent difference (%) between $C_{RNA,\ ACL}$ and $C_{RNA,\ expected}$ (Fig. S10B). *CV* (%) is the ratio of the standard deviation to the mean $C_{RNA,\ ACL}$, and is used to describe variability *within* replicates of the same M gene RNA concentration. We found non-linear relationships (Fig. S10A, black dotted lines with R^2 values < 0.28) between *CV* (%) and M gene RNA concentration at all sampled cycle numbers. We also calculated the percent difference (%) to describe variability *between* $C_{RNA,\ ACL}$ and $C_{RNA,\ expected}$ using Eq. S14:

$$\%\ difference = \frac{log(C_{RNA,ACL}) - log(C_{RNA,expected})}{log(C_{RNA,expected})} * 100 \quad \textbf{(Eq. S14)}$$

Again, we found a non-linear relationship across RNA concentrations (Fig. S10B, black dotted lines with R^2 values < 0.36). We note that variability in conversion with ACL standard curves is comparable to that of the *Ct* method, using *CV* (%) (Fig. S10A, *Ct*, R^2 = 0.02) and percent difference (%) (Fig. S10B, *Ct*, R^2 = 0.25).

(R5) Validating ACL Standard Curves for Multiple Template Sequences. ACL standard curves were also validated for two additional template sequences, the human MYCN gene DNA and soybean Lectin endogene DNA. These were compared to the previously discussed M gene dataset (Fig. S11A). The M gene dataset consisted of 21 RNA concentrations (Table S2, 3.28×10^3 to 2.62×10^9 copies/µL) used in previous positive control and validation experiments. The human MYCN gene and soybean Lectin endogene DNA datasets were taken from previously published studies on the evaluation of qPCR curve analysis methods (14) and enhanced analysis of qPCR data using a variable efficiency model (7). The MYCN dataset consisted of four DNA concentrations (ranging from 1.88×10^9 to 1.88×10^3 copies/µL) quantified with a SYBR Green qPCR assay. The Lectin dataset consisted of five dilutions (ranging from 6.4×10^0 to 4.0×10^3

copies/μL), also quantified with a SYBR Green qPCR assay. For each of the three template sequences, we constructed an ACL library from reference curves of 10^7 copies/μL for the M gene, 1875 copies/μL for the MYCN gene, and 800 copies/μL for the Lectin endogene. The standard curve used in the M gene dataset was constructed from ACL cycle 20. The standard curves used in MYCN and Lectin datasets were from ACL cycles 27 and 29, respectively.

ACL conversions of F_N to template concentration ($C_{RNA/DNA,\ ACL}$) for the IAV M gene RNA (Fig. S11A, green dots), human MYCN DNA (Fig. S11A, blue squares), and soybean Lectin endogene DNA (Fig. S11A, red circles) datasets were plotted as a linear regression against the known template concentrations of each dilution ($C_{RNA/DNA,\ expected}$). For each dilution, there were 3, 94, or 18 technical replicates for the M gene, MYCN, or soybean datasets, respectively. There was a clear linear relationship between $C_{RNA/DNA,\ ACL}$ and $C_{RNA/DNA,\ expected}$ for all datasets (Fig. S11A, $R^2 >$ 0.9640). Only 2 of the 30 total $C_{RNA/DNA,\ ACL}$ values (from all datasets) fell outside a 2-fold change (Fig. S11A, black dotted lines) from $C_{RNA/DNA,\ expected}$. Of the 28 $C_{RNA/DNA,\ ACL}$ values (93% of the data) that fell within a 2-fold change, there was a maximum fold change of 1.7.

To further demonstrate the improved fit of reference amplification curves with the SCF-E model (Eq. S5) as opposed to the SCF model (8,13), the M gene dataset was also converted using a reference curve fit with the SCF model (Fig. S11A, gray dots). There was a weaker linear relationship between $C_{RNA,\ ACL}$ and $C_{RNA,\ expected}$ (R^2 = 0.8487), with 8 of the 21 $C_{RNA/DNA,\ ACL}$ values falling outside of the 2-fold from $C_{RNA,\ expected}$.

Furthermore, we compared ACL standard curves to *Ct* standard curves for all three template sequences. Conversion with *Ct* standard curves showed a similar linear relationship between $C_{RNA/DNA,\ ACL}$ and $C_{RNA/DNA,\ expected}$ for all datasets (Fig. S11B, $R^2 > 0.9912$). This time, all measured $C_{RNA/DNA,\ ACL}$ values fell within a 2-fold change of the expected mean with a maximum fold change of 1.9.

443 **(R6) Validating dqPCR for Multiple PCR Cycle Numbers.** To examine the contribution of error
444 within individual PCR cycle numbers to overall error in the dqPCR distributions (Fig. 3B), we
445 performed an Analysis of Variance (ANOVA) test (Fig. S15) on random samples of 100 drops
446 from each M gene distribution in Fig. 3B. For the 17.1 M gene cpd group (sampled at N = 22 to
447 25), measured cpd differed by cycle number (p-value <0.05) and accounted for 8% of the total
448 error in pooled measurements. For the 171 M gene cpd group (sampled at N = 19 to 22),
449 measured cpd differed by cycle number (p-value <0.05) and accounted for 59% of the total error
450 in pooled measurements. For the 1710 M gene cpd group (sampled at N = 6 to 22), M gene
451 measured cpd differed by cycle number (p-value <0.05) and accounted for 7% of the total error in
452 pooled measurements. The higher error in the 171 cpd was an outlier, perhaps due to
453 experimental variability from cycle 19, as error was reduced to 18% when that cycle was
454 excluded; nevertheless, the other concentrations showed relatively lower error. To address this
455 concern, we further examined the reported variability between cycle numbers using another
456 metric, the linear mixed effects model (LME, Fig. S16). The percent difference (%) (Eq. S14)
457 between the measured and expected M gene cpd was calculated at each sampled cycle number.
458 The LME used cycle number as the fixed effect variable and expected M gene cpd as the random
459 effect variable. Both variables were log-transformed during the model fit, as replication cycles
460 correspond to RNA copy numbers on a log scale. The fixed effects coefficients described a
461 negative relationship between the slope and percent difference (%) from the expected M gene
462 cpd of -0.85% per cycle number (p-value <0.05, SE = 0.01). While the p-value indicates
463 significance, the magnitude of the relationship between cycle and converted cpd is small (~1%),
464 reinforcing the validity of pooling data converted at different cycle numbers together. Furthermore,
465 when we pool cycle numbers in our control experiment (Fig. 3B), distributions were centered
466 around the expected mean (Table S6) and had a linear relationship between measured and
467 expected cpd, that also fell within a 2-fold change from the expected mean (Fig. S17).
468
469 **(R7) Extracting Individual Distributions from Mixed Droplet Detection Data.** To use dqPCR
470 for measuring IAV burst size across a population of single-cell infections, which we expected to

be heterogeneous, we validated that our method allows for isolation of individual distributions from a sample containing multiple M gene RNA concentrations. This was done by applying a Gaussian mixture model (GMM) (Table S7) to the distributions shown in Fig. 3E.

We began by fitting three log-normal distributions to the three individual concentrations of M gene cpd that make up the mixed drop distribution in Fig. 3E. We let x_i be M gene cpd such that, at x_1 = 1.71 × 10^1, x_2 = 1.71 × 10^2, or x_3 = 1.71 × 10^3cpd. The dqPCR model attributes measurement y_i to each drop. Because of measurement noise, x_i and y_i are generally not equal. Mathematically, we represent the relationship between the measurement y_i of a drop containing x_i cpd, using Eq. S15:

$$\mathbf{log_{10}}(\boldsymbol{y_i}) = \mathbf{log_{10}}(\boldsymbol{x_i}) + \boldsymbol{E_i} \quad \textbf{(Eq. S15)}$$

Here, E_i is the measurement noise. Inspired by the shape of the distribution in Fig. 3E, we assumed that E_i follows a Gaussian distribution with mean B_i (representing measurement bias) and variance σ_i^2 (representing the amount of measurement noise). Each Gaussian has a different weight corresponding to the probability that a random drop in the mixture has x_i copies. Hence, to fit data in Fig. 3E, we needed to estimate these weights w_i along with the B_i and σ_i^2 values. The maximum likelihood estimates are represented in Table S7. In general, our model describes the measured distribution well, as seen from the fit in Fig. 3E.

Results from Table S7 suggest that the measurement bias $\hat{B}_i$ decreases linearly with increases in the cpd on a log scale. We therefore used a linear function to describe the relationship in Eq. S16:

$$\boldsymbol{B(x)} \approx \mathbf{-0.06 \times log_{10}}(\boldsymbol{x}) + \mathbf{0.20} \quad \textbf{(Eq. S16)}$$

Here, the constants in Eq. S16 are estimated from the values reported in Table S7.

For the standard deviation of the measurement noise, $\hat{\sigma}_t$, we notice that the change is non-monotonic with respect to increases in cpd. We therefore fitted a quadratic curve to describe their relationship in Eq. S17:

$$\sigma(x) = 0.63 - 0.47x + 0.11x^2 \quad \textbf{(Eq. S17)}$$

Here, the constants are also estimated from the values reported in Table S7.

(R8) Filtering of Drops Containing Cell Lysate from Burst Size Distributions. To ensure that burst size measurements are solely from extracellular viral particles, and do not include intracellular viral RNA from cell lysate, we developed a multiplexed dqPCR assay that simultaneously detects cellular β-actin mRNA and IAV M gene RNA. β-actin was chosen as a cellular indicator due to its abundance as a structural protein, its highly conserved sequence between cell types, and its linear relationship to cell concentration (Fig. S18). For each burst size replicate experiment (Table S8), two reference amplification curves were generated using 10^2 cpd of M gene and β-actin template controls (Table S1) in 50 µm drops. As expected, drops containing high β-actin mRNA also exhibited high concentrations of M gene RNA (Fig. S19). These drops were presumed to contain non-packaged, intracellular M gene RNA. To remove these drops from the burst size data, we used one of two thresholds. The threshold was set at either the top 12.5% of drops containing β-actin, corresponding to the microfluidic chip split ratio, or above a β-actin concentration of 2×10^3 cpd, corresponding to the amount of β-actin released from a lysed cell (slope of Figure S18, $\approx 2 \times 10^3$ β-actin copies per cell). The threshold which removed the most drops was chosen to ensure that drops containing lysed cells were excluded from the data.

527 **(R9) Oseltamivir (OST) treatment of IAV infected cells during drop infections.**
528 To demonstrate that burst size measurements include only RNA from extracellular viral particles,
529 we measured IAV production from single cells treated with oseltamivir acid (OST), the active
530 metabolite of the antiviral drug oseltamivir. OST functions as a neuraminidase inhibitor,
531 preventing the cleavage of assembled IAV particles from the host cell membrane. MDCK cells, a
532 commonly used cell line for producing high-titer viral stocks, were infected with IAV H1N1 were
533 encapsulated into 100 µm diameter microfluidic drops using methods for 'IAV Infection' and
534 'Droplet Encapsulation of Infected Cells' described in the main text. During encapsulation,
535 infected cells were either suspended in standard droplet infection media (+H1N1 -OST) or in
536 droplet infection media treated with 10 µM concentration of oseltamivir (+H1N1 +OST). Mock
537 infected cells encapsulated in standard infection media (-H1N1 -OST) were used as a negative
538 control. M gene abundance (copies/µL) was measured using a bulk RT-qPCR assay from the
539 supernatant of broken drops sampled at 0 and 18 hpi. The 0 hpi bulk samples were frozen
540 overnight at -80 °C while the 18 hpi bulk samples were incubated overnight at 37 °C. Preparation
541 of drop infections for bulk RT-qPCR involved freezing a 400 µL sample of drops at -80 °C for 30
542 mins to break the emulsion. 300 µL of the broken emulsion was collected for further processing.
543 Collected bulk and drop infections were clarified by centrifugation at 500 × *g* for 5 min, and the
544 resulting supernatant was sampled for RT-qPCR targeting the IAV M gene. IAV production in
545 drops had a mean of 1.930×10^3 M gene copies/µl at 18 hpi, across three technical replicates
546 (Figure S21). In the presence of oseltamivir, IAV production was detected in only one of three
547 technical replicates, and measured 4.004 copies/µl. This was a significant decrease according to
548 a two-sample Student's t-test ($p << 0.05$). These results support that burst size measurements
549 include only IAV particles released from the host cell membrane and are not contaminated by
550 exosomal RNA or other forms of extracellular contamination, as oseltamivir prevents viral budding
551 and release.
552
553 **(R10) Modeling IAV Burst Size Distributions from Droplet RNA Concentration.** The
554 measured burst sizes distributions in Fig. 4B are compounded by two underlying distributions, the

555 viral burst size distribution and the measurement noise. To infer the viral burst size distribution,
556 we first estimated the measurement noise by fitting log-normal distributions to fluorescence data
557 for a mixed sample of three known M gene RNA concentrations in drops (Fig. 3E, 1.71×10^1,
558 1.71×10^2, or 1.71×10^3 cpd). We found that the standard deviation of the distributions ranged
559 from 0.15 - 0.30 on a $\log_{10}$ scale (Table S7) and used this to describe the measurement noise.
560
561 We then fit a model incorporating the measurement noise and three parametric distributions,
562 Poisson, negative-binomial and log-normal (Fig. S22), against the burst size measurements in
563 Fig. 4B. First, we simulated the viral burst size from the assumed distribution. For each simulated
564 value of burst size x, we generated simulated measurement noise assuming a log-normal
565 distribution with a mean defined by the bias function $B(x)$ and a standard deviation defined by
566 $\sigma(x)$. Then, we computed a density function (kernel density with a Gaussian kernel) for the
567 resulting distribution and used this function to compute the log-likelihood of observations. We
568 estimated the parameters of each distribution and for each of the H3N2 or H1N1 populations by
569 maximizing the likelihood of observations reported in Fig. 4B. To reduce the computational cost,
570 we first obtained rough parameter estimates by simulating 10,000 values from each potential
571 burst size distribution. The ensuing estimates were then refined by repeating the above
572 procedure using 100,000 values while constraining parameter values to a domain closer to the
573 rough estimates.
574
575 Comparing the goodness-of-fit of the three modeled distributions to the burst size data using
576 Akaike Information Criterion (AIC) scores, we determined that the burst size measurements were
577 best described by a negative-binomial distribution (Table S10 and Fig. 4D). The estimated mean
578 burst sizes of these negative binomial distributions were 709 and 358 for H3N2 and H1N1,
579 respectively. The shape parameter of each negative binomial distribution was estimated to be
580 0.48 and 0.49 for H3N2 and H1N1, respectively. Note that the negative-binomial distribution is
581 considered highly dispersed when the shape parameter is less than 1. Therefore, there exists
582 large heterogeneity in the burst size distribution for both H1N1 and H3N2.

583

584

585

Supplementary figures

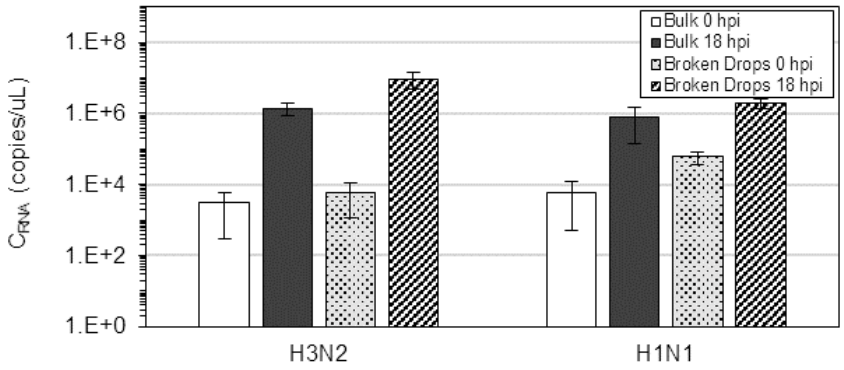
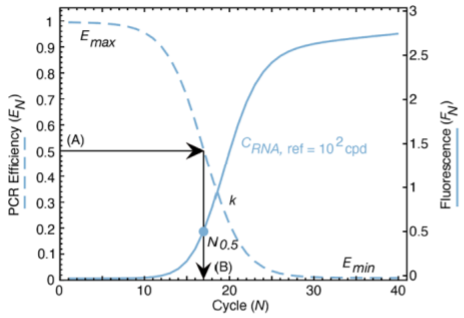


Figure S1. Comparison of IAV M Gene Abundance During Bulk and Drop Infections. IAV strains (**H1N1** and H3N2) were separately used to infect A549 cells under bulk cell culture and single cell encapsulation in 100 µm diameter microfluidic drops. M gene abundance (copies/µL) was measured using a bulk RT-qPCR assay from the supernatant of both bulk and drop infections at 0 and 18 hpi. Experimental details can be found in **SI R1.** Each bar represents the pooled data from three replicate experiments. Error bars represent one standard deviation.

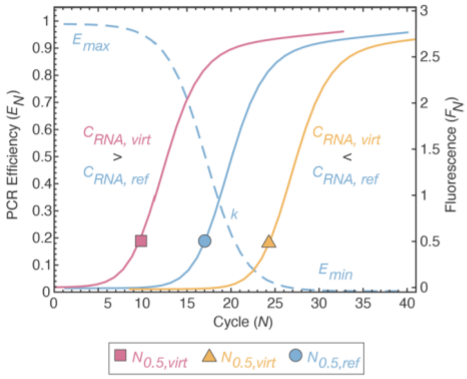


(Eq. S3) $$E_N = \frac{E_{max}}{1 + e^{\left(\frac{N - N_{0.5}}{k}\right)}} + E_{min}$$

(Eq. S4) $$F_{N+1} = F_N(1 + E_N)$$

(Eq. S5) $$F_{N+1} = F_N\left(1 + \frac{E_{max}}{1 + e^{\left(\frac{N - N_{0.5}}{k}\right)}} + E_{min}\right)$$

Figure S2. Sigmoidal Curve Fitting using PCR Efficiency (SCF-E). A four-parametric sigmoid function was used to model decreasing PCR efficiency (E_N) at each cycle number (N) (Eq. S3). This model was fitted to discontinuous fluorescence measurements (F_N) of a known RNA template concentration ($C_{RNA,\ ref}$) to generate a continuous reference amplification curve (Eq. S4 and S5). Reference curves used in burst size experiments were generated with $C_{RNA,\ ref}$ = 1.71 × 10^2 RNA (cpd). In the SCF-E model, E_N is the PCR efficiency at cycle N; E_{max} is the maximum PCR efficiency at cycle number 1 and ranges from of 0.9 to 1; E_{min} is the minimum PCR efficiency at cycle number 40 and ranges from 0 to 0.1; $N_{0.5}$ is the cycle number at which E_N equals 0.5; and k is the shape parameter of the curve. As k increases, the shape of the curve flattens, and as k decreases, the shape of the curve becomes steeper. Thirty values of each of the four efficiency parameters (E_{max}, E_{min}, $N_{0.5}$, *and* k) were used to test 8.1×10^5 (30^4) curve fits to experimental fluorescence measurements. The set of efficiency parameters yielding the best fit were used to construct the SCF-E reference curve (solid blue curve).

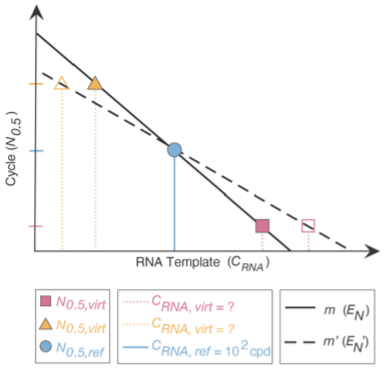


(Eq. S7)

$$F_{N+1} = F_N\left(1 + \frac{E_{max}}{1 + e^{\left(\frac{N - N_{0.5,virtual}}{k}\right)}} + E_{min}\right)$$

610

611 **Figure S3. Translating a Single SCF-E Reference Curve into 1000 Virtual Amplification**
612 **Curves.** ACLs consist of virtual amplification curves that share the same E_{max}, E_{min}, and k
613 parameters as the reference curve but have unique $N_{0.5}$ values. This is done by substituting $N_{0.5,}$
614 $_{ref}$ from Eq. S5 with $N_{0.5,\ virt}$ in Eq. S7. We input 1000 evenly spaced $N_{0.5,\ virt}$ values between cycle
615 numbers N = 1 to 40, resulting in 1000 virtual amplification curves, each associated with a unique
616 theoretical RNA concentration ($C_{RNA,\ virt}$). Virtual curves with $N_{0.5,\ virt} < N_{0.5,\ ref}$ (leftmost pink solid
617 curve) correspond to higher RNA concentrations compared to the reference curve, whereas
618 curves with $N_{0.5,\ virt} > N_{0.5,\ ref}$ (rightmost yellow solid curve) represent lower RNA concentrations
619 than the reference (middle blue solid curve). The efficiency curve for reference and virtual curves
620 is illustrated as a dashed blue line.
621
622
623
624
625
626
627
628
629



(Eq. S8)

$$N_{0.5} = m * log_{10}(C_{RNA}) + b$$

(Eq. S10)

$$log_{10}(C_{RNA,virt}) = log_{10}(C_{RNA,ref}) + \frac{N_{0.5,virt} - N_{0.5,ref}}{m}$$

(Eq. S13)

$$m = \frac{1}{log_{10}(E_N + 1)}$$

630

631 **Figure S4. Calculating RNA Template Concentration of the Virtual Amplification Curves.**
632 We relate the known RNA concentration of the reference curve ($C_{RNA,\ ref}$) to the unknown RNA
633 concentrations of the virtual curves ($C_{RNA,\ virt}$) based on their position on the x-axis (cycle number,
634 N). In Eq. S10, we define a relationship between the known cycle numbers where PCR efficiency
635 is equal to 0.5 for the reference ($N_{0.5,\ ref}$) and virtual ($N_{0.5,\ virt}$) curves and their respective RNA
636 concentrations (C_{RNA}). By rearranging Eq. S10, we can calculate the unknown $C_{RNA,\ virt}$ using Eq.
637 S12. To perform the calculation, the slope of the line (m) in Eq. S13 is determined using a PCR
638 efficiency constant (E_N). As indicated by Eq. S13, an increase in E_N leads to a decrease in m
639 (dotted black line), expanding the range of $C_{RNA,\ virt}$ in the ACL (from solid triangle to open triangle,
640 and from solid square to open square). This relationship is visually illustrated here with Eq. S8.
641

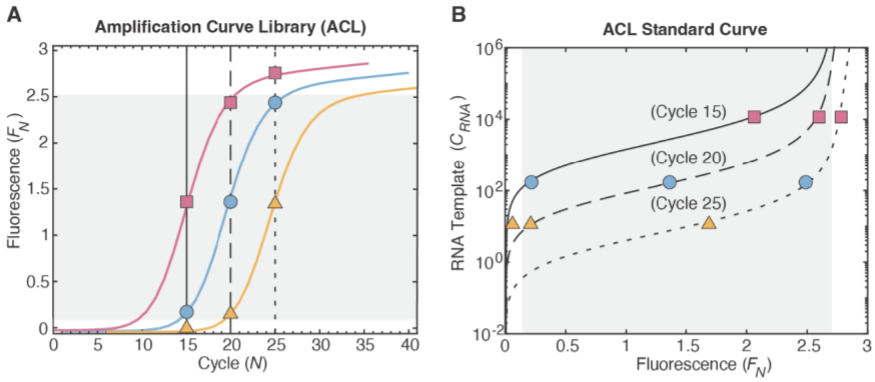
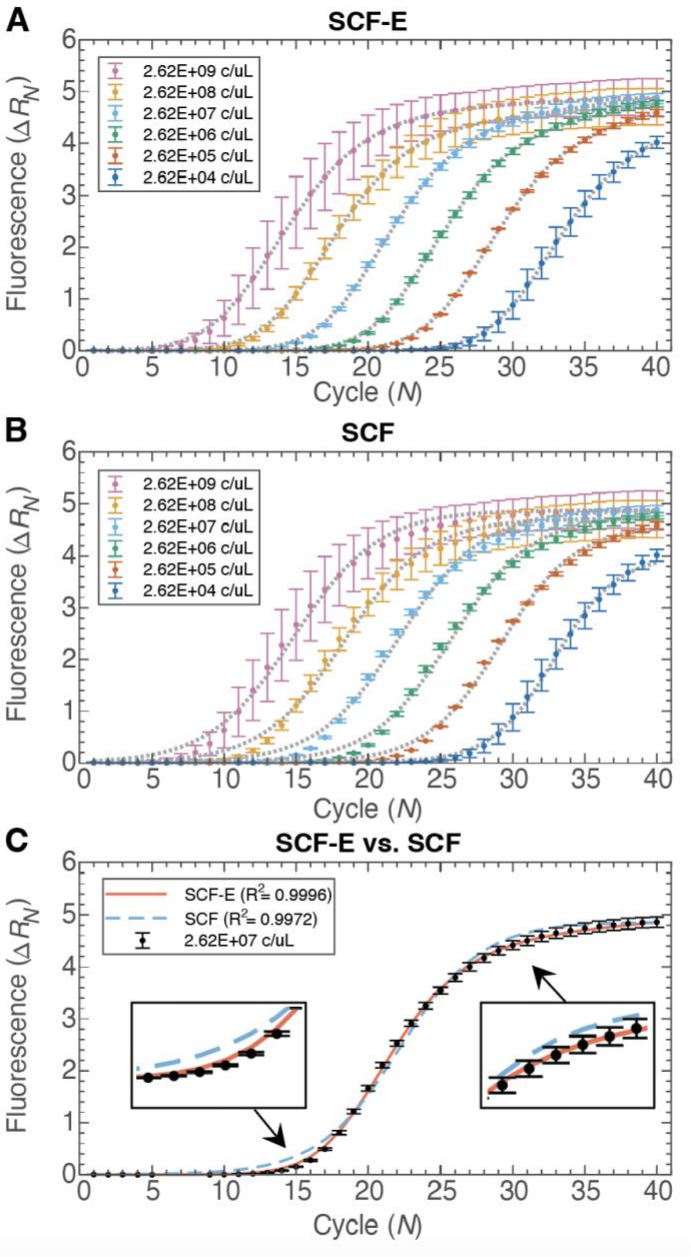


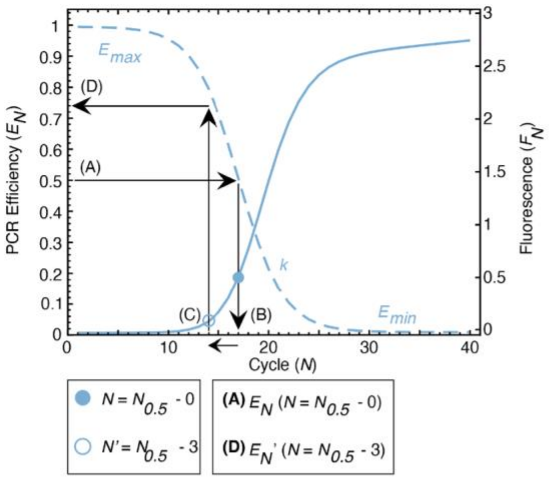
Figure S5. Constructing a Standard Curve from the ACL. (A) Construction of an amplification curve library (ACL) (Fig. S3). Three $C_{RNA,\ virt}$ within the ACL are displayed here (pink, blue, yellow), with the leftmost pink curve representing a higher $C_{RNA,\ virt}$ compared to the rightmost yellow curve. **(B)** ACL standard curves (solid line, dashed line, dotted line) are constructed from the F_N of $C_{RNA,\ virt}$ (pink, blue, yellow) at a particular cycle number (15, 20, 25). The cycle number chosen for the ACL standard curve corresponds to the cycle number at which unknown RNA concentrations are sampled. In this method, only virtual curves in the exponential or linear phase at the selected cycle number can be included in the standard curve (grey shaded regions in **A** and **B**). This is because the early amplification and plateau regions of multiple virtual curves may overlap at a single cycle number. To ensure that only useable regions are selected for the standard curve, we establish threshold values based on the fluorescence of the reference curve. The upper threshold is set at 1 standard deviation below the F_N at cycle 40, and the lower threshold is set at the 99th percentile of the F_N at cycle 1. For example, as amplification proceeds for a particular $C_{RNA,\ virt}$ (pink curve) in **A**, the F_N increases and intersects cycles 15, 20, and 25 at the three points (squares). However, only two of those points are within the valid range of the ACL standard curves in **B.**



660

661 **(Continued on Next Page)**

662 **Figure S6. Comparison of SCF-E and SCF Methods of Fitting PCR Amplification Curves.**
663 Fluorescence intensities (ΔR_N) of bulk RT-qPCR amplification of six concentrations of IAV M gene
664 RNA (Table S2, 10-fold dilutions from $2.62 \cdot 10^4$ to $2.62 \cdot 10^9$ copies/µL) were measured at N = 1
665 to 40. The concentration of each dilution decreases from left to right when viewing the
666 amplification curves. **(A)** Amplification curves generated with the SCF-E model (dotted gray
667 curves) had an average fit of R^2 = 0.9996 across all RNA concentrations (Table S3). **(B)**
668 Amplification curves generated with the SCF model (dotted gray curves) had an average fit of R^2
669 = 0.9975 across all RNA concentrations (Table S3). **(C)** For a single M gene RNA concentration
670 ($2.62 \cdot 10^7$ copies/µL), the SCF-E model provided a better fit (R^2 = 0.9996) than SCF (R^2 =
671 0.9972). The difference in fit is most noticeable at the lower ($\Delta R_N < 1$) and higher ($\Delta R_N > 4$)
672 fluorescence values, representing the exponential and plateau phases of PCR amplification
673 (magnified insets). All error bars represent one standard deviation from the average fluorescence
674 measured across three technical replicates.



(Eq. S3)

$$E_N = \frac{E_{max}}{1 + e^{\left(\frac{N - N_{0.5}}{k}\right)}} + E_{min}$$

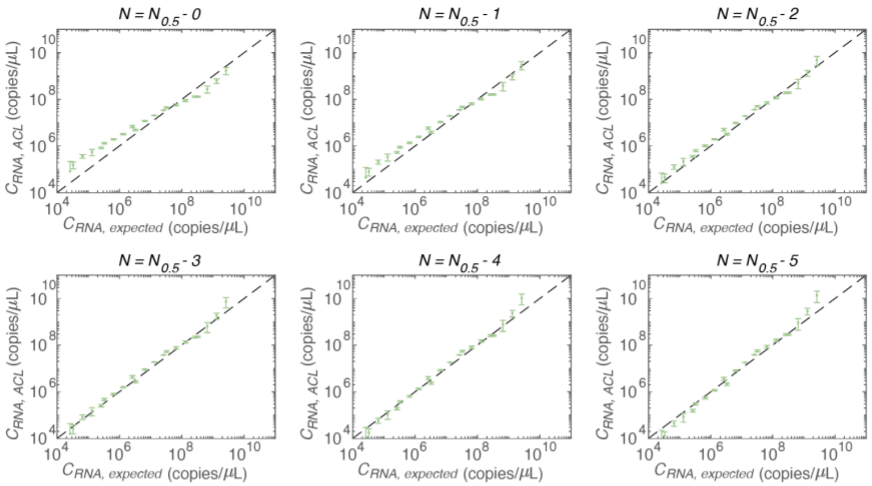
(D) $E_N'\,(N = N_{0.5} - 3)$

$$E_N' = \frac{E_{max}}{1 + e^{\left(\frac{-3}{k}\right)}} + E_{min}$$

675

676 **Figure S7. Setting a PCR Efficiency Constant (E_N) for the ACL.** To calculate E_N at $N = N_{0.5} -$
677 *3*, we determined $N_{0.5}$ of the SCF-E reference curve, which corresponds to N where $E_N = 0.5$ **(A**
678 **to B)**. $N_{0.5} - 3$ represents three cycles prior to $N_{0.5}$ **(C)**. We recorded the E_N at $N_{0.5} - 3$ **(D)**. When N
679 $= N_{0.5} - 3$ is inserted into Eq. S3, the term $N_{0.5}$ cancels out, resulting in a constant value of -3 for
680 $(N - N_{0.5})$. The E_N calculated in this step is used in Eq. S13 to determine m.

681



682

683 **Figure S8. Linear regression of measured to expected RNA concentrations from Table S4.**
684 We compared cycle numbers (N) at which PCR efficiency (E_N, Eq. S3) of the reference curve is
685 used to calculate the constant m (Eq. S13), for assigning RNA template concentrations to virtual
686 curves in the ACL. We tested the cycle number at which PCR efficiency is equal to 0.5 ($N_{0.5}$) and
687 the five cycles preceding it ($N = N_{0.5}$ *-1 to* $N = N_{0.5}$ *-5*), to construct six ACLs for comparison. From
688 each ACL, a standard curve was constructed at PCR cycle number 20. We evaluated the
689 dynamic range of each standard curve in Table S4. The ACL standard curve with the widest
690 dynamic range and the highest degree of accuracy (R^2) between measured ($C_{RNA,\ ACL}$) and
691 expected ($C_{RNA,\ expected}$) M gene RNA concentrations was built using E_N at $N = N_{0.5} - 3$. For this
692 ACL, the dynamic range of the standard curve was 4.70 $\log_{10}$ RNA copies/μL (Table S4, dynamic
693 range of 2.62×10^4 to 1.31×10^9 copies/μL) with $R^2 = 0.990$ between $C_{RNA,\ ACL}$ and $C_{RNA,\ expected}$.
694 Therefore, the E_N at $N = N_{0.5} - 3$ was used to construct all amplification curve libraries in this
695 work.
696

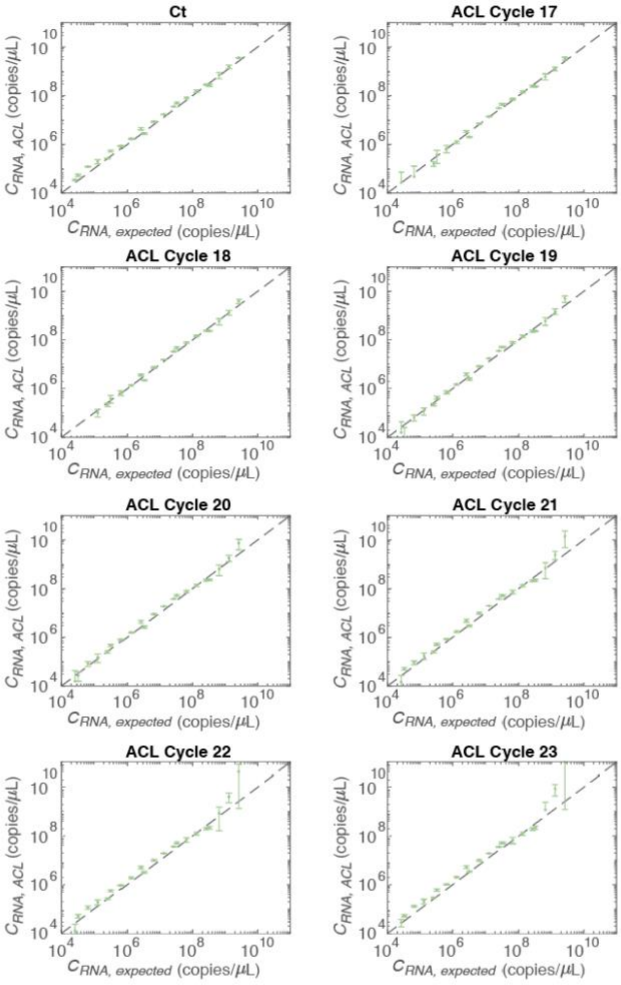
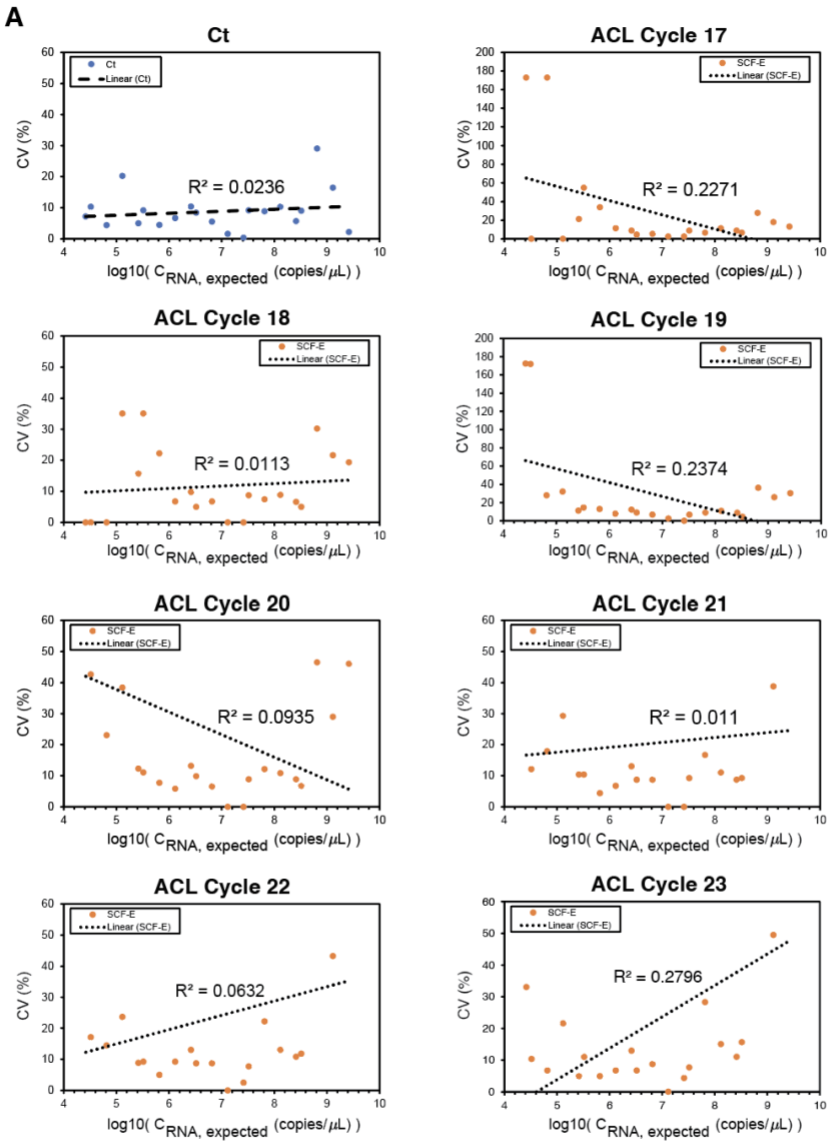


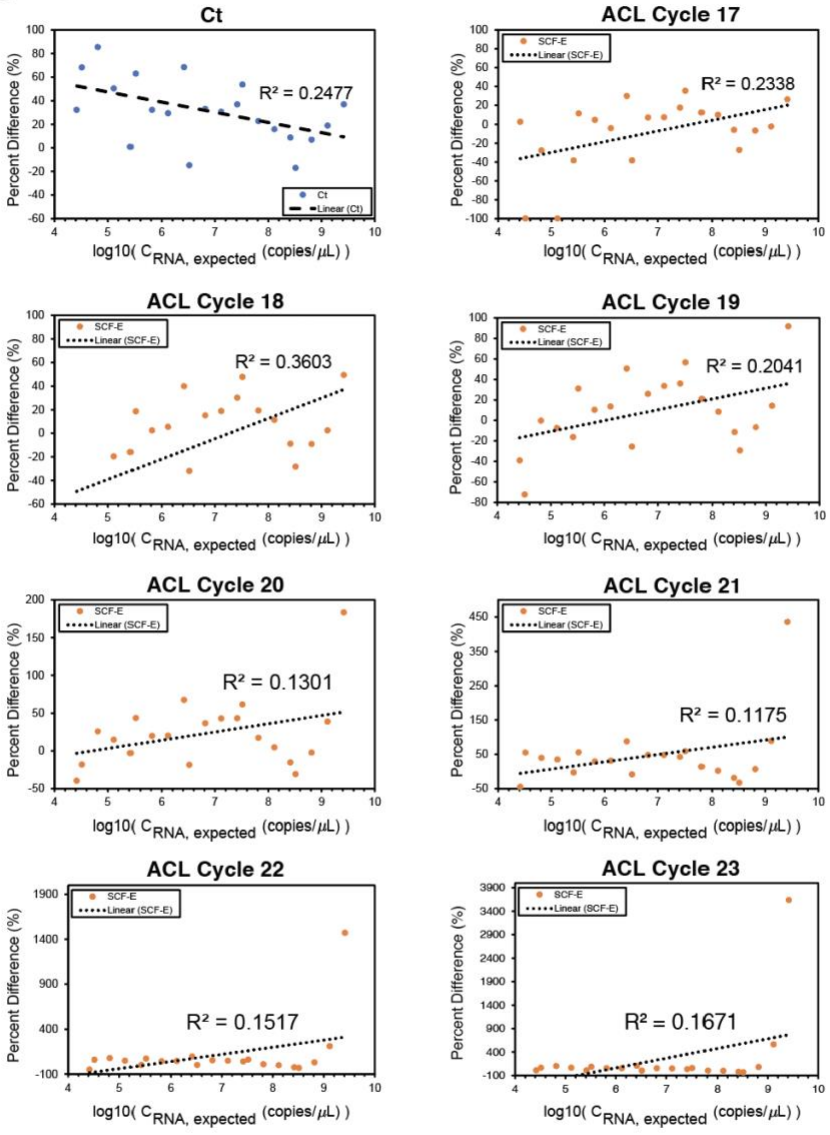
Figure S9. Linear regression of measured to expected RNA concentrations from Table S5. We compared the dynamic range of ACL standard curves from different cycle numbers (N = 17 to 23) to the Ct standard curve. Standard curves were used to convert fluorescence to template concentration for a known dilution series of IAV M gene RNA (Table S2, three replicates of 21 concentrations between 2.62 × 10^4 to 2.62 × 10^9 copies/uL). Standard curve dynamic range and degree of accuracy (R^2) between measured ($C_{RNA, ACL}$) and expected ($C_{RNA, expected}$) M gene RNA concentrations were evaluated in Table S5.



705

706 **(Continued on Next Page)**

B



(Continued on Next Page)

710 **Figure S10. Variability in M gene RNA Concentrations Measured with ACL Standard**
711 **Curves by PCR Cycle Number.** We examined variability between standard curves obtained from
712 different ACL cycle numbers using two measures: **(A)** coefficient of variation (*CV* %) and **(B)**
713 percent difference (%). Standard curves were constructed from ACL cycle numbers (N = 17 to 23,
714 individual plots) and used to convert fluorescence to RNA concentration ($C_{RNA, ACL}$) across 23 IAV
715 M gene RNA concentrations (Table S2). **(A)** *CV* (%) is the ratio of the standard deviation to the
716 mean $C_{RNA, ACL}$, and it quantifies the variability *within* replicates of the same M gene RNA
717 concentration. A linear regression (dotted line) was conducted to compare *CV* (%) across M gene
718 RNA concentrations. The results revealed a weak linear relationship ($R^2 < 0.28$) for all cycle
719 numbers. **(B)** Percent difference (%) measures the variability *between* $C_{RNA, ACL}$ and the expected
720 RNA concentration ($C_{RNA, expected}$). Similarly, a weak relationship ($R^2 < 0.36$) was observed
721 between the starting RNA concentration and percent difference (%) for all cycle numbers. To
722 assess the variability in ACL standard curves, a comparison was made with *Ct* standard curves
723 (A and B, first upper left plot), for *CV* (%) and percent difference (%). The analysis indicated no
724 linear response for *CV* (%) ($R^2 = 0.0236$) and a weak linear response for percent difference (%)
725 ($R^2 = 0.2477$) across the M gene RNA concentrations.
726

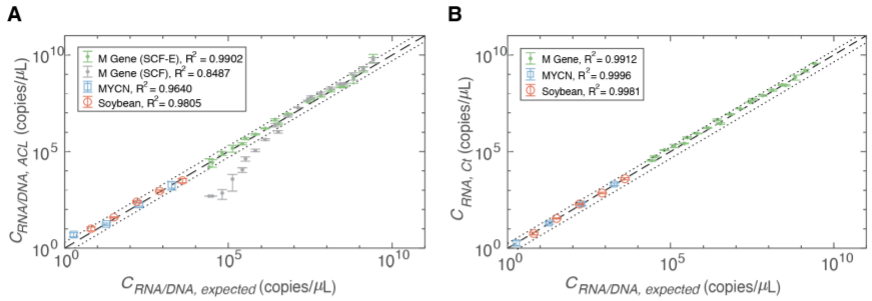


Figure S11: Validating ACL Standard Curves for Multiple Template Sequences. Standard curves constructed using the **(A)** ACL method or **(B)** Ct method were used to convert fluorescence to template concentration ($C_{RNA/DNA,\ ACL}$) for a known dilution series of three different template sequences: IAV M gene RNA (green/gray dots, ranging from 2.62×10^4 to 2.92×10^9 copies/μL), human MYCN gene DNA (blue squares, ranging from 1.88×10^0 to 1.88×10^3 copies/μL), and soybean Lectin endogene DNA (red circles, ranging from 6.4×10^0 to 4.0×10^3 copies/μL). Each dilution of the templates had varying numbers of technical replicates: 3 for the M gene, 94 for the MYCN, and 18 for the Lectin dataset. The M gene ACL was built with a 10^7 copies/μL reference curve fitted with either the SCF-E (green dots) or SCF (grey dots) model. The standard curve from the M gene ACL was constructed at a single cycle number (N = 20). The MYCN libraries were built with a 1,875 copies/μL reference curve, and the Soybean ACL was built with an 800 copies/μL reference curve, both fitted using the SCF-E model. The standard curves from these ACLs were constructed at N = 27 (MYCN) and N = 29 (Lectin). An R^2 value quantified the linear response between measured $C_{RNA/DNA,\ ACL}$ to expected concentrations ($C_{RNA/DNA,\ expected}$) across template concentrations. Dotted lines indicate a 2-fold change from $C_{RNA/DNA,\ expected}$, and error bars represent one standard deviation from the mean.

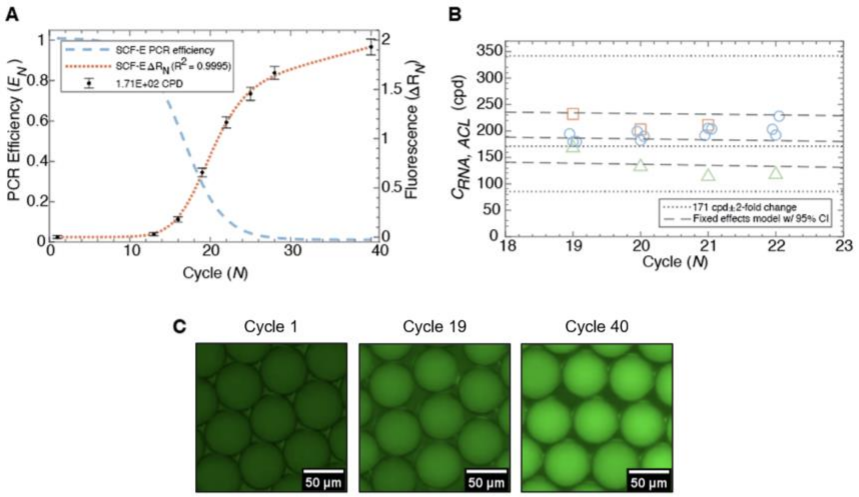
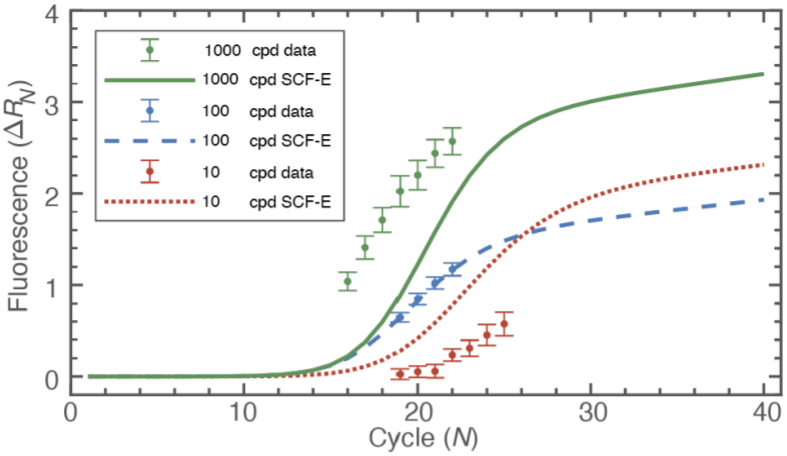


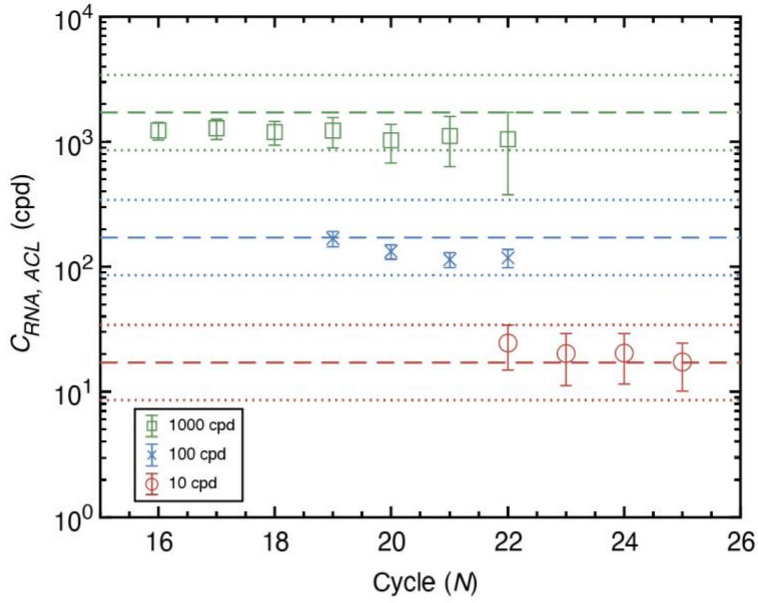
Figure S12. Droplet Quantitative PCR (dqPCR) of the IAV M gene. (A) SCF-E reference amplification curve generated for 1.71 x 10^2 cpd of IAV M gene RNA amplified in 50 µm drops. Drop fluorescence intensity (ΔR_N) was detected at PCR cycle numbers N = 1, 13, 16, 19, 22, 25, 28, 40 and displayed as a mean (black dots) and one standard deviation (error bars) (≈1000 or more drops per cycle number). The estimated PCR efficiency (E_N) curve (dashed blue line) used during SCF-E produced a well-fitting reference amplification curve (dotted red line, R^2 = 0.9995). **(B)** dqPCR was validated for a single concentration of M gene RNA (1.71x10^2 cpd) across three biological replicates. The samples were collected at various cycle numbers: rep 1 (red squares) at N = 19 to 21 and rep 2 (blue circles) and rep 3 (green triangles) at N = 19 to 22. Drop fluorescence was converted to M gene RNA concentration ($C_{RNA, ACL}$) within a 2-fold change of the expected mean (dotted lines) for all cycle numbers. Dashed lines indicate a fixed effects model with upper and lower bounds set as the 95% confidence intervals. **(C)** Representative epifluorescence images of 50 µm diameter drops at PCR cycle numbers 1, 19, and 40.



760

761 **Figure S13. SCF-E Reference Amplification Curves used in Figure 3B.** Three known
762 concentrations of M gene RNA, 1.71×10^1 cpd (red dots), 1.71×10^2 cpd (blue dots), and $1.71 \times$
763 10^3 cpd (green dots), were amplified in 50 µm drops. Drop fluorescence (ΔR_N) was detected at
764 multiple PCR cycle numbers (Table S6, note that cycles 19-21 in the 10 cpd group were below
765 background and not included in the table) and converted to M gene cpd using dqPCR (Fig. 3B).
766 For each experiment run on different days, a new SCF-E reference curve was generated from the
767 amplification of 1.71×10^2 cpd (red dotted line, blue dashed line, green solid line).

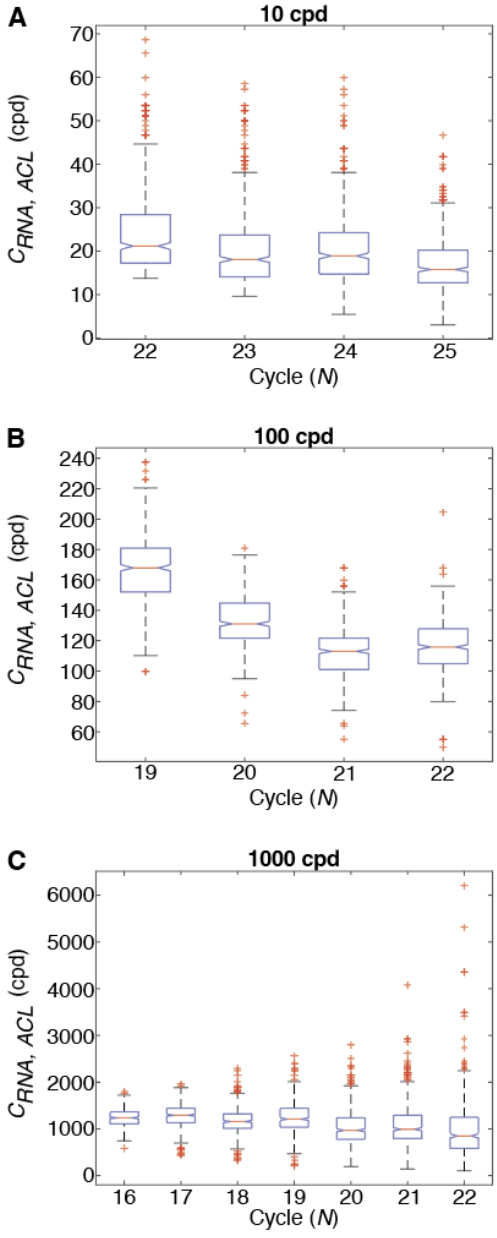
768



769

770 **Figure S14. Using dqPCR to Quantify Three IAV M gene RNA Concentrations at Multiple**
771 **Cycle Numbers.** Three known IAV M gene RNA concentrations, 1.17×10^1 cpd (10 cpd, red
772 circles), 1.17×10^2 cpd (100 cpd, blue crosses), and 1.17×10^3 cpd (1000 cpd, green squares)
773 were amplified in 50 µm drops. Drop fluorescence (ΔR_N) was detected at multiple PCR cycle
774 numbers (Table S6) and converted to M gene cpd ($C_{RNA,\ ACL}$) using dqPCR. $C_{RNA,\ ACL}$
775 measurements were pooled together and displayed as a single distribution in Fig. 3B. Dashed
776 lines indicate the expected M gene cpd, and dotted lines represent a 2-fold change from the
777 expected mean.

778



779

780

(Continued on Next Page)

781 **Figure S15. Analysis of Variance (ANOVA) for dqPCR of Three IAV M Gene RNA**
782 **Concentrations.** Three known IAV M gene RNA concentrations, 1.17×10^1 cpd (10 cpd plot),
783 1.17×10^2 cpd (100 cpd plot), and 1.17×10^3 cpd (1000 cpd plot), were amplified in 50 µm drops.
784 Drop fluorescence (ΔR_N) was detected at multiple PCR cycle numbers (Table S6) and converted
785 to M gene cpd ($C_{RNA,\,ACI}$) using the dqPCR model. $C_{RNA,\,ACI}$ measurements were pooled together
786 and displayed as a single distribution in Fig. 3B. To examine the contribution of error within
787 individual PCR cycle numbers to overall error in the pooled distributions, we performed an
788 Analysis of Variance (ANOVA) test on a random sample of 100 drops from each group. **(A)** The
789 10 cpd group, quantified at PCR cycle numbers 22 – 25, had mean $C_{RNA,\,ACI}$ values that differed
790 by cycle number (ANOVA p-value <0.05). This variability due to cycle number accounted for 8%
791 of the total error in the pooled distribution. **(B)** The 100 cpd group, quantified at PCR cycle
792 numbers 19-22, had mean $C_{RNA,\,ACI}$ values that differed by cycle number (ANOVA p-value <0.05).
793 This variability due to cycle number accounted for 59% of the total error in the pooled distribution.
794 **(C)** The 1000 cpd group, quantified at PCR cycle numbers 16-22, had mean $C_{RNA,\,ACI}$ values that
795 differed by cycle number (ANOVA p-value <0.05). This variability due to cycle number accounted
796 for 7% of the total error in the pooled distribution.
797

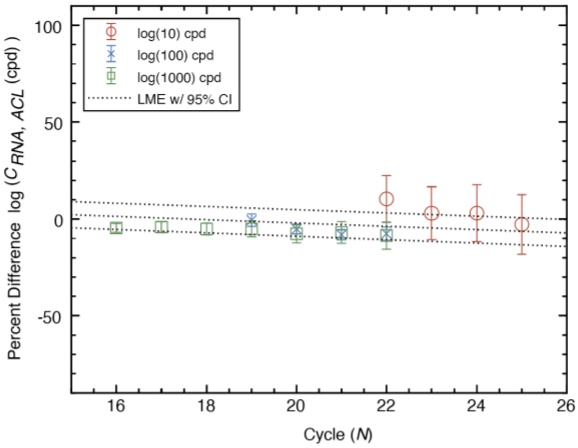
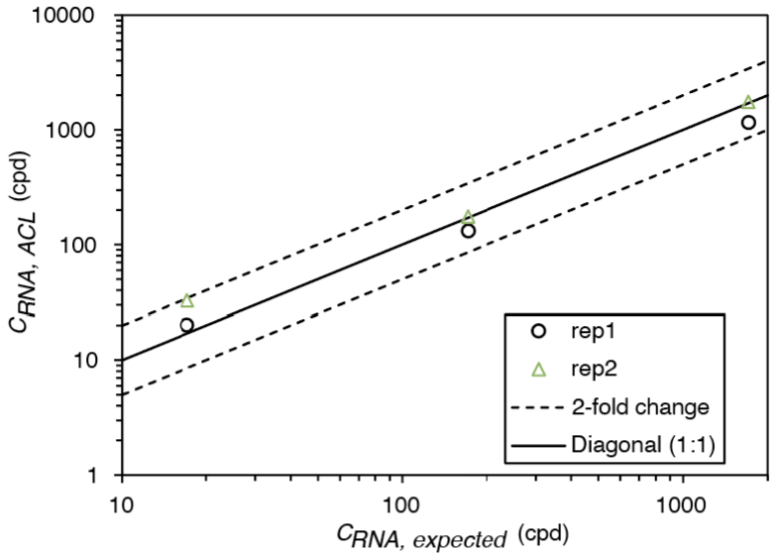
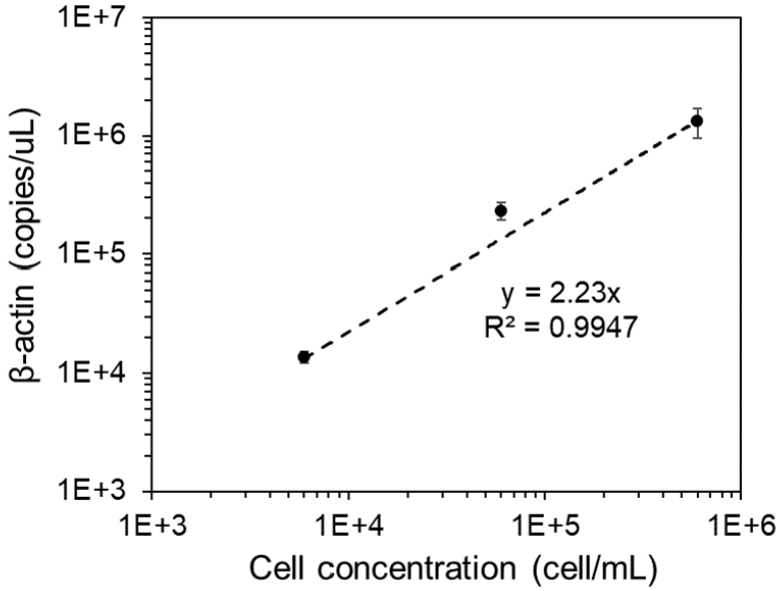


Figure S16. Linear Mixed Effects Model (LME) for dqPCR of Three IAV M Gene RNA Concentrations. Three known IAV M gene RNA concentrations, 1.17×10^1 cpd (10 cpd, red circle), 1.17×10^2 cpd (100 cpd, blue cross), and 1.17×10^3 cpd (1000 cpd, green triangle), were amplified in 50 µm drops. Drop fluorescence (ΔR_N) was detected at multiple PCR cycle numbers (Table S6) and converted to M gene cpd ($C_{RNA,\ ACL}$). $C_{RNA,\ ACL}$ measurements were pooled together and displayed as a single distribution in Fig. 3B. To further examine the contribution of error within individual PCR cycle numbers to overall error in the pooled distributions, we calculated the percent difference (%) (Eq. S14) between the measured $C_{RNA,\ ACL}$ values and their expected M gene cpd ($C_{RNA,\ expected}$), across cycle numbers. This analysis informed the validity of pooling conversions measured from multiple PCR cycles in our burst size experiments. Percent differences were compared using a linear mixed effects model (LME) with a 95% confidence interval (CI) (dotted lines). In the LME, we set cycle as the fixed effect variable and $C_{RNA,\ expect}$ as the random effect variable. Both variables were log-transformed during the model fit, as replication cycles correspond to RNA copy numbers on a log scale. The fixed effects coefficients described a negative relationship between the slope and percent difference from the expected RNA concentration of -0.85% per cycle number (p-value <0.05, SE = 0.01). While the p-value indicates significance, the magnitude of the relationship is small (~1%/cycle), maintaining that it is valid to pool dqPCR data from different cycle numbers.



817

818 **Figure S17. Using dqPCR to Quantify IAV M gene RNA Concentration, Pooled Cycle**
819 **Numbers.** Three known IAV M gene RNA concentrations, 1.17×10^1 cpd, 1.17×10^2 cpd, $1.17 \times$
820 10^3 cpd, were amplified in 50 μm drops. Drop fluorescence (ΔR_N) was detected at multiple PCR
821 cycle numbers (Table S6) and converted to M gene cpd ($C_{RNA,\ ACL}$) using dqPCR. $C_{RNA,\ ACL}$
822 measurements were pooled together and plotted against the expected M gene cpd ($C_{RNA,\ expected}$).
823 Two biological replicates (black circle and green triangle) are shown. The solid line represents a
824 1:1 relationship between $C_{RNA,\ ACL}$ and $C_{RNA,\ expected}$, dashed lines indicate a 2-fold change from the
825 expected mean.
826



827

828 **Figure S18. β-actin mRNA abundance increases with cell concentration.** A bulk RT-qPCR
829 assay was performed to amplify β-actin mRNA from lysed A549 cells after 24 hr of incubation.
830 There was a strong linear relationship between β-actin and cell concentration (dashed black line,
831 R^2 = 0.995), validating its use as a reliable marker for the detection of intracellular lysate in drops
832 during our burst size experiments.

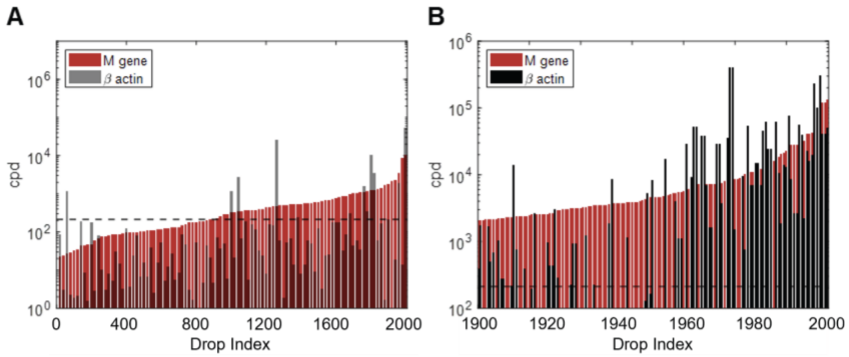


Figure S19. Multiplexed dqPCR for Detection of IAV M gene and Cellular β-actin. (A) Both M gene RNA (red) and β-actin mRNA (black) were detected by multiplexed dqPCR from H1N1 infection in drops (n=2000). Data is presented in ascending order of RNA concentration, prior to β-actin filtering (indicated by the dashed black line). **(B)** Showing a zoomed-in region of (A), with drop indexes 1900 to 2000.

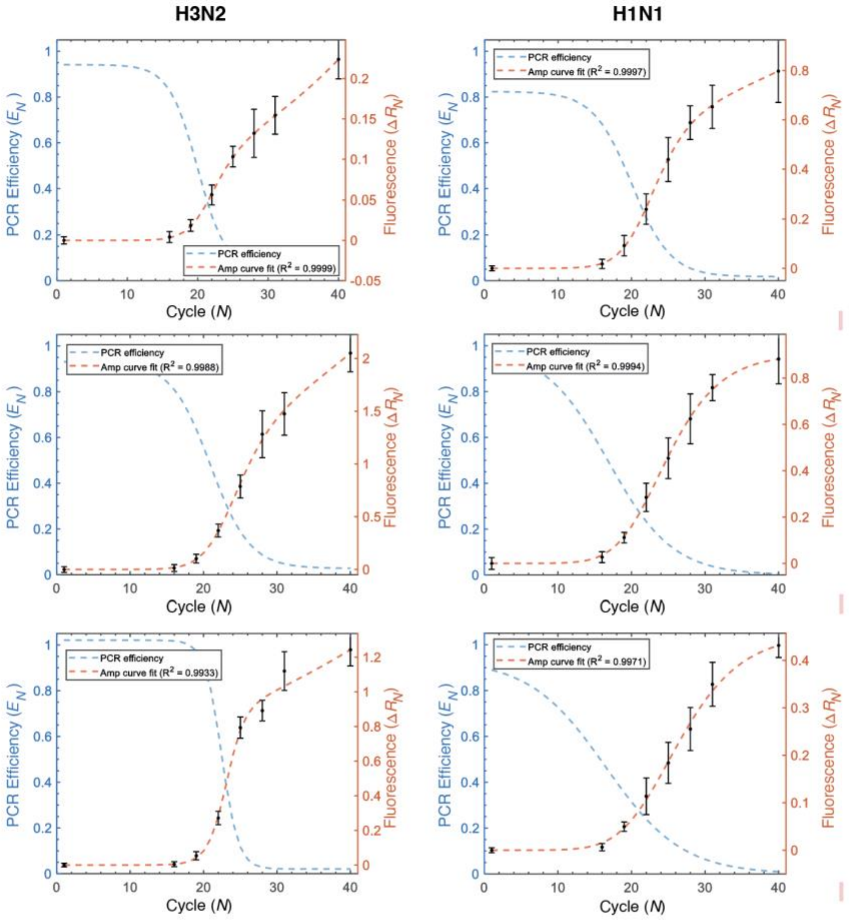


Figure S20. SCF-E Reference Curves for Burst Size Replicate Experiments. For each burst size replicate experiment (Table S8), a new reference amplification curve was constructed for dqPCR (Table S9). To generate these curves, a known concentration of M gene RNA (1.71 x 10^2 cpd) was amplified in 100 µm drops. Drop fluorescence intensity (ΔR_N) was measured at PCR cycle numbers N = 1, 16, 19, 22. 25, 28, 31 and 40. The ΔR_N of drops from each cycle number are shown as the mean (black dots) with error bars representing one standard deviation. SCF-E was used to create a continuous reference amplification curve (orange dashed line) from discontinuous ΔR_N measurements, using an estimate for the PCR efficiency (blue dashed line).

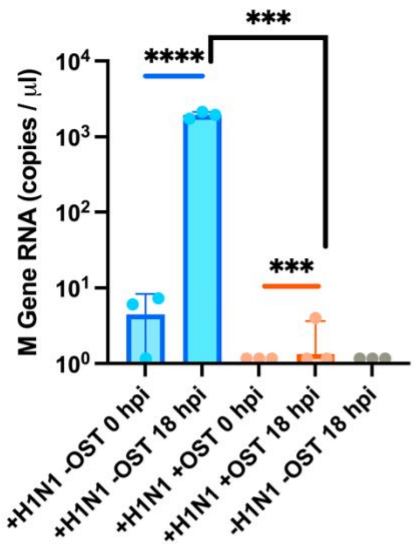


Figure S21. Oseltamivir (OST) treatment of IAV infected cells during drop infections. MDCK cells infected with IAV H1N1 were encapsulated into 100 µm diameter microfluidic drops. During encapsulation, infected cells were either suspended in standard droplet infection media (+H1N1 -OST) or in droplet infection media treated with 10 µM concentration of oseltamivir (+H1N1 +OST). Mock infected cells encapsulated in standard infection media (-H1N1 –OST) were used as a negative control. M gene abundance (copies / µL) was measured using a bulk RT-qPCR assay from the supernatant of broken drops sampled at 0 and 18 hpi. Each bar represents the pooled data from three technical replicates. Significance stars (*** $p < 0.001$ and **** $p < 0.0001$) obtained by a two-sample Student's t-test.

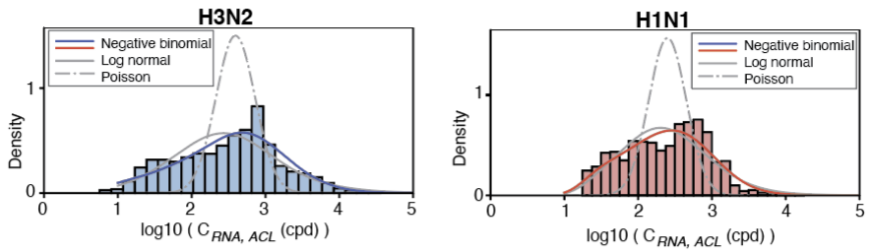


Figure S22. Best Fit Model to Estimate IAV Burst Size Distributions from Table S10.

We used a simulation-based approach to estimate the IAV burst distributions shown in Figure 4D. We considered three distribution models: lognormal (solid gray curve), Poisson (dashed gray curve), and negative-binomial distribution (solid blue or red curve, corresponding to H3N2 and H1N1 distributions, respectively). The parameters for each distribution are provided in Table S10. Based on the analysis, burst sizes were found to best fit a negative binomial distribution.

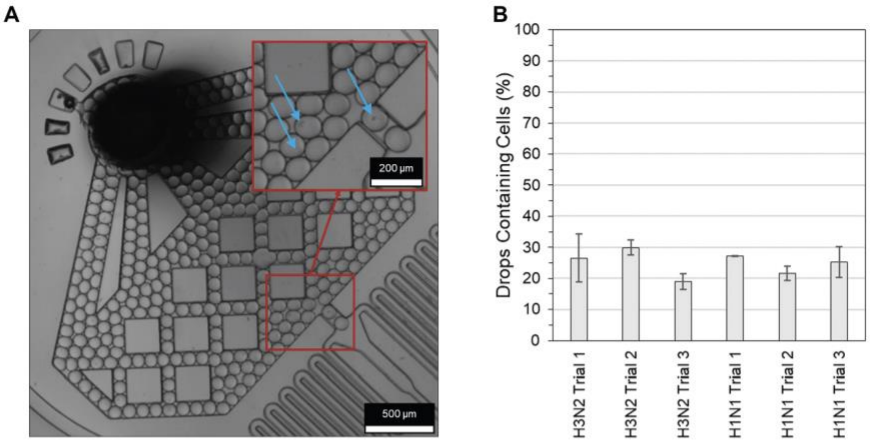


Figure S23. Cell Loading During Burst Size Replicate Experiments. **(A)** Images of 100 µm drops re-injected into a microfluidic device were captured at the beginning and end of each burst size replicate experiment to monitor cell loading. Drops containing a single cell are indicated by blue arrows. **(B)** Across all H3N2 and H1N1 replicate experiments, cell loading was determined by counting the percentage of drops containing cells (mean = 25% ± 4%). Cell loading in drops was assumed to follow a Poisson distribution. For a population of drops with 25% containing cells, the estimated Poisson mean (λ) is 0.29 cells/drop. In this case, ≈75% of drops are empty, ≈21.5% of drops contain one cell, ≈3% contain two cells, and ≈0.5% of drops contain three or more cells. The bar graph displays the drop counts for the replicates from left to right, which are 753, 679, 538, 719, 774, and 582.

879

50

880 ## Supplementary tables

881 **Table S1. RT-qPCR Template and Primer Sequences**.

	Sequence (5' to 3')
M gene dsDNA template control	GTCTAATACGACTCACTATAG*GACCAATCCTGTCACCTCTGAC***TG** **CAGTCCTCGCTCACTGGGCACG**TGCTTCATCGCGAACTGCTTCG CGGATGCCATCGTCATGGCCACGAGGATATGTAAGAGT*TAGACG ATTTGTCCAGAATGCCCT*
M gene forward primer	GACCRATCCTGTCACCTCTGAC
M gene reverse primer	AGGGCATTCTGGACAAATCGTCTA
M gene Taqman probe	/FAM/TGCAGTCCTCGCTCACTGGGCACG/BHQ1/
β-actin dsDNA template control	GTGGATGATGATATCGCCGCGCTCGTCGTCGACAACGGCTCCG GCATGTGCAAGGCCGGCTTCGCGGGCGACGATGCCCCCCGGG CCGTCTTCCCCTCCATCGTGGGGCGCCCCAGGCACCAGGGCGT GATGGTGGGCATGGGTCAGAAGGATTCCTATGTGGGCGACGAG GCCCAGAGCAAGAGAGGCATCCTCACCCTGAAGTACCCCATCGA GCACGGCATCGTCACCAACTGGGACGACATGGAGAAAATCTGGC ACCACACCTTCTACAATGAGCTGCGTGTGCCTCCCGAGGAGCAC CCCGTGCTGCTGACCGAGGCCCCCCTGAACCCCAAGGCCAACC GCGAGAAGATGACCCAGATCATGTTTGAGACCTTCAACACCCCA GCCATGTACGTTGCTATCCAGGCTGTGCTATCCCTGTACGCCTCT GGCCGTACCACTGGCATCGTGATGGACTCCGGTGACGGGGTCA CCCACACTGTGCCCATCTACGAGGGGTATGCCCTCCCCCATGCC ATCCTGCGTCTGGACCTGGCTGGCCGGGACCTGACTGACTACCT CATGAAGATCCTCACCGAGCGCGGCTACAGCTTCACCACCACGG CCGAGCGGGAAATCGTGCGTGACATTAAGGAGAAGCTGTGCTAC GTCGCCCTGGACTTCGAGCAAGAGATGGCCACGGCTGCTTCCA GCTCCTCCCTGGAGAAGAGCTACGAGCTGCCTGACGGCCAGGT

	CATCACCATTGGCAATGAGCGGTTCCGCTGCCCTGAGGCACTCT TCCAGCCTTCCTTCCTGGGCATGGAGTCCTGTGGCATCCACGAA ACTACCTTCAACTCCATCATGAAGTGTGACGTGGACATCCGCAAA GACCTGTACGCCAACACAGTGCTGTCTGGCGGCACCACCATGTA CCCTGGCATTGCCGACAGGATGCAGAAGGAGATCACTGCCCTG GCACCCAGCACAATGAAGATCAAGATCATTGCTCCTCCTGAGCG CAAGTACTCC*GTGTGGATCGGCGGCTCCATC*CTGGCCTCGCTGT CC**ACCTTCCAGCAGATGTGGATC**AG*CAAGCAGGAGTATGACGA* *GTC*CGGCCCCTCCATCGTCCACCGCAAATGCTTCTAG
β-actin forward primer	GTGTGGATCGGCGGCTCCATC
β-actin reverse primer	GACTCGTCATACTCCTGCTTG
β-actin Taqman probe	/Cy5/ACCTTCCAGCAGATGTGGATC/BHQ2/

All sequences are listed in the 5' to 3' direction. The M gene template control was constructed as linear dsDNA containing a T7 promoter (underlined), forward and reverse primer binding sites (*italicized*), a probe binding site (**bold**), and a partial M gene sequence. The M gene dsDNA was *in vitro* transcribed prior to use as a positive control. The M gene TaqMan probe carried a FAM fluorophore with a BHQ1 quencher. The β-actin template control was constructed as a dsDNA plasmid containing forward and reverse primer binding sites (*italicized*), and a probe binding site (**bold**), and a partial β-actin mRNA sequence. The β-actin TaqMan probe carried a Cy5 fluorophore with a BHQ2 quencher.

Table S2. IAV M Gene RNA Concentrations Amplified by Bulk RT-qPCR and Used to Validate dqPCR.

M gene RNA concentration (copies/µL) of positive controls used in bulk RT-qPCR assays
2.62×10^4
3.28×10^4
6.55×10^4
1.31×10^5
2.62×10^5
3.28×10^5
6.55×10^5
1.31×10^6
2.62×10^6
3.28×10^6
6.55×10^6
1.31×10^7
2.62×10^7
3.28×10^7
6.55×10^7
1.31×10^8
2.62×10^8
3.28×10^8
6.55×10^8
1.31×10^9
2.62×10^9

895 **Table S3. SCF-E and SCF Curve Fitting Parameters for Figure S6.**

M gene RNA concentration (copies/μL)	E_{min} (unitless)	E_{max} (unitless)	$N_{0.5}$ (cycle)	k (cycle)	R^2 (SCF-E)	R^2 (SCF)
2.62×10^9	0.002	1.000	9.1	3.31	0.9994	0.9965
2.62×10^8	0.003	1.000	13.1	3.03	0.9995	0.9967
2.62×10^7	0.005	1.000	17.1	2.97	0.9996	0.9972
2.62×10^6	0.009	1.000	20.8	2.83	0.9997	0.9976
2.62×10^5	0.010	0.993	24.4	2.90	0.9998	0.9984
2.62×10^4	0.010	1.000	28.9	2.83	0.9997	0.9988
Average	**0.007**	**0.999**	**18.9**	**2.98**	**0.9996**	**0.9975**

896

Table S4. Dynamic Range of ACL Standard Curves Built with Different E_N Constants.

PCR Cycle (N)	E_N	m	$C_{RNA,\ min}$ (copies/µL)	$C_{RNA,\ max}$ (copies/µL)	Dynamic Range ($\log_{10}$(copies/µL))	R^2
$N_{0.5}$ - 0	0.514	5.55	3.28×10^6	1.31×10^8	1.60	0.886
$N_{0.5}$ - 1	0.597	4.92	3.28×10^6	2.62×10^8	1.90	0.960
$N_{0.5}$ - 2	0.675	4.46	2.62×10^4	2.62×10^9	5.00	0.986
$N_{0.5}$ - 3	0.745	4.14	2.62×10^4	1.31×10^9	4.70	0.990
$N_{0.5}$ - 4	0.805	3.90	3.28×10^4	1.31×10^9	4.60	0.986
$N_{0.5}$ - 5	0.853	3.73	6.55×10^4	6.55×10^8	4.00	0.979

We compared cycle numbers (N) at which PCR efficiency (E_N) of the reference curve is used to calculate the constant m (and E_N). This constant determines the RNA template concentration assigned to virtual curves in the ACL. E_N is calculated using Eq. S3 and m is calculated using Eq. S13. We used the cycle number at which PCR efficiency is equal to 0.5 ($N_{0.5}$) and the five cycles preceding it ($N = N_{0.5}$ *-1 to* $N = N_{0.5}$ *-5*), to construct six ACLs for comparison. From each ACL, a standard curve was generated at PCR cycle number 20. This standard curve was then used to convert fluorescence (F_N) to template concentration (C_{RNA}) for a known dilution series of M gene RNA (Table S2, three replicates of 21 concentrations between 2.62×10^4 to 2.62×10^9 copies/µL). The minimum ($C_{RNA,\ min}$) and maximum ($C_{RNA,\ max}$) concentrations in the dilution series, whose measured RNA concentration fell within a 2-fold change of the expected concentrations, marked the dynamic range of each standard curve. The R^2 value quantifies the linear regression (Fig. S8) between the measured ($C_{RNA,\ ACL}$) and expected ($C_{RNA,\ expected}$) RNA concentrations. The ACL built with E_N at $N = N_{0.5} - 3$ produced a cycle 20 standard curve with the highest dynamic range and the greatest degree of accuracy (R^2). Therefore, an offset of -3 cycles was used to construct all ACLs in this work.

914 **Table S5. Dynamic Range of ACL Standard Curves from Different ACL Cycle Numbers.**

Method	Standard Curve PCR Cycle (*N*)	$C_{RNA,\ min}$ (copies/µL)	$C_{RNA,\ max}$ (copies/µL)	Dynamic Range (log_{10}(copies/µL))	R^2
Ct	NA	2.62×10^4	2.62×10^9	5.00	0.991
ACL	17	2.62×10^5	2.62×10^9	4.00	0.812
ACL	18	1.32×10^5	2.62×10^9	4.30	0.800
ACL	19	6.55×10^4	2.62×10^9	4.60	0.989
ACL	20	2.62×10^4	1.31×10^9	4.70	0.990
ACL	21	2.62×10^4	1.31×10^9	4.70	0.981
ACL	22	3.28×10^5	6.55×10^8	4.30	0.962
ACL	23	6.55×10^4	6.55×10^8	4.00	0.938

915

916 We compared the dynamic range of standard curves created using the amplification curve library
917 (ACL) method to that of the cycle threshold (*Ct*) method. Standard curves were used to convert
918 fluorescence to template concentration for a known dilution series of IAV M gene RNA (Table S2,
919 three replicates of 21 concentrations between 2.62×10^4 to 2.62×10^9 copies/µL). In the ACL
920 method, multiple standard curves were constructed from individual cycle numbers (N = 17 to 23).
921 On the other hand, in the *Ct* method, a single standard curve is built from multiple *Ct* values, with
922 one *Ct* value corresponding to each concentration of M gene RNA. The dynamic range of each
923 standard curve was determined by identifying the minimum ($C_{RNA,\ min}$) and maximum ($C_{RNA,\ max}$)
924 concentrations within the dilution series that yielded measured RNA concentration falling within a
925 2-fold change of the expected concentrations. The R^2 value quantifies the linear response (Fig.
926 S9) between the measured ($C_{RNA,\ ACL}$) and expected ($C_{RNA,\ expected}$) RNA concentrations.
927

928 **Table S6. Using dqPCR to Quantify IAV M Gene RNA Concentrations at Multiple Cycle**
929 **Numbers.**

Group	M gene RNA concentration (cpd)	PCR Cycle (*N*)	Mean $C_{RNA,\ ACL}$ (cpd)	Standard Deviation (cpd)	*CV* (%)	Number of Drops (n)
10^1	17.1	22	24.49371	9.605813	39.21746	734
	17.1	23	20.18378	9.014579	44.66249	1472
	17.1	24	20.33817	8.819321	43.3634	1282
	17.1	25	17.26811	7.173752	41.54335	1025
	17.1	**Pooled**	**20.2664**	**8.955352**	**44.18817**	**4513**
10^2	171	19	167.8969	23.34381	13.90366	3479
	171	20	132.3851	16.98913	12.83312	4268
	171	21	113.7438	15.53087	13.65426	5540
	171	22	117.5996	19.49319	16.57589	2612
	171	**Pooled**	**131.2311**	**27.83475**	**21.21048**	**15899**
10^3	1710	16	1228.969	196.8349	16.01627	10107
	1710	17	1279.32	237.4656	18.56186	10443
	1710	18	1194.993	259.0783	21.68032	9186
	1710	19	1226.291	334.1291	27.24713	8743
	1710	20	1025.578	350.1932	34.14592	9569
	1710	21	1111.12	479.4927	43.15398	9335
	1710	22	1046.932	669.8305	63.98033	7878
	1710	**Pooled**	**1163.231**	**390.8254**	**33.59827**	**65261**

930
931 **(Continued on Next Page)**
932

933 Three known IAV M gene RNA concentrations, 1.71×10^1 cpd (10^1), 1.71×10^2 cpd (10^2), and
934 1.71×10^3 cpd (10^3), were amplified in 50 µm drops. Drop fluorescence (ΔR_N) was detected at
935 multiple PCR cycle numbers and converted to M gene cpd ($C_{RNA,\ ACL}$) using dqPCR. $C_{RNA,\ ACL}$
936 measurements from different cycle numbers were pooled together and presented as a single
937 distribution in Figure 3B. To assess the variability between individual cycle numbers and the
938 pooled distribution, we analyzed the standard deviation of the mean $C_{RNA,\ ACL}$ and the coefficient
939 of variation (*CV* %) across the number of sampled drops (n) at each N.
940

Table S7. Maximum Likelihood Estimates for a Gaussian Mixture Model (GMM) Used to Extract Individual Distributions from Mixed Droplet Detection Data.

x_i (cpd)	Weights ($\widehat{w_i}$)	Bias ($\widehat{B_i}$)	Standard Deviation ($\hat{\sigma}_i$)
1.71×10^1	0.18	0.11	0.22
1.71×10^2	0.35	0.07	0.15
1.71×10^3	0.47	-0.01	0.30

To use dqPCR for measuring IAV burst size in a heterogeneous population of single-cell infections, we validated the method's ability to isolate individual distributions from a sample containing multiple M gene RNA concentrations. We applied a Gaussian mixture model (GMM) on dqPCR data obtained from three known IAV M gene RNA concentrations: 1.17×10^1 cpd, 1.17×10^2 cpd, and 1.17×10^3 cpd, mixed together in a single sample (data shown in Figure 3E). Model assumptions are described in **SI R7**. In the model, x_i represents the number of M gene RNA copies (cpd). Each Gaussian has a different weight (w_i) corresponding to the probability that a random drop in the mixture has x_i copies, a predicted mean (B_i, representing measurement bias) and variance (σ_i^2, representing measurement noise).

954 **Table S8. Burst Size Replicate Experiments.**

Strain	Replicate	Collected Drops (n)	Filtered Drops (n)	Mean Burst Size (viruses/cell)	Burst Size IQR (viruses/cell)	Gini coefficient (*G*)
H1N1	1	2141	1873	4.9×10^2	6.3×10^1 to 6.9×10^2	0.623
H1N1	2	2840	1932	4.8×10^2	1.1×10^2 to 6.8×10^2	0.520
H1N1	3	2036	1780	4.8×10^2	9.9×10^1 to 5.0×10^2	0.601
H1N1	**Pooled**	**7017**	**5585**	**4.8×10^2**	**8.8×10^1 to 6.3×10^2**	**0.586**
H3N2	1	4524	3955	1.6×10^3	1.1×10^2 to 1.6×10^3	0.724
H3N2	2	4288	3751	6.3×10^2	1.1×10^2 to 8.1×10^2	0.581
H3N2	3	2193	1914	3.8×10^2	5.3×10^1 to 5.8×10^2	0.518
H3N2	**Pooled**	**11005**	**9620**	**1.0×10^3**	**9.3×10^1 to 8.9×10^2**	**0.705**

955

956 IAV burst sizes were measured from thousands of individual drops sampled at cycle numbers *N* =
957 16, 19, 22, 25, and 28 in three biological replicates. The measurements were used to determine
958 the number of IAV M gene RNA and cellular β-actin mRNA (cpd). Drops containing high β-actin
959 cpd (Fig. S19) were assumed to contain non-packaged, intracellular, M gene RNA and were
960 subsequently filtered from the final burst size distributions (Filtered Drops, *SI Appendix* R8). We
961 present the average measured burst size for each replicate, along with the interquartile range
962 (IQR) that represents the lower 25th and upper 75th percentile of the distributions. The
963 heterogeneity of each distribution was examined using the Gini coefficient (*G*). If $G = 0$, the
964 distribution is completely homogeneous, meaning that each cell produces the same number of
965 viruses. Conversely, if $G = 1$, the distribution is completely heterogeneous, where one cell
966 produces all the viruses. Thus, a higher *G* corresponds to a more heterogeneous distribution.
967

968 **Table S9. dqPCR Parameters Used in Burst Size Replicate Experiments.**

Strain	Replicate	E_{min} (unitless)	E_{max} (unitless)	$N_{0.5}$ (cycle)	k (cycle)	$N_{0.5}$ - 3	E_N	m
H1N1	1	0.0172	0.807	20.0	2.62	17.0	0.629	4.72
H1N1	2	0	0.993	16.8	4.38	13.8	0.660	4.54
H1N1	3	0	0.934	16.5	5.17	13.5	0.599	4.91
H3N2	1	0.0414	0.907	20.0	2.03	17.0	0.779	4.00
H3N2	2	0.0276	0.903	20.9	2.66	17.9	0.710	4.29
H3N2	3	0.0207	1	22.4	1.14	19.4	0.954	3.44

969

970 The parameters used to create reference standard curves using SCF-E (Fig. S20, orange dotted
971 curves) for each burst size experiment.

972

973 **Table S10. Best Fit Model to Estimate IAV Burst Size in Fig. 4D.**

Strain	Burst size distribution	Parameter estimates		Mean Burst Size (cpd)	AIC
H3N2	Poisson	mean = 359.4 cpd		359	81541.5
H3N2	**Negative Binomial**	**shape = 0.48**	**prob = 0.9993**	**709**	**19938.2**
H3N2	Log-Normal	mean = 5.57 cpd	std= 1.70	1120	20302.4
H1N1	Poisson	mean = 222.4 cpd		222	25394.9
H1N1	**Negative Binomial**	**shape = 0.49**	**prob = 0.9986**	**358**	**8689.0**
H1N1	Log-Normal	mean = 5.24 cpd	std= 1.44	534	9044.2

974

975 We used a simulation-based approach to estimate the IAV burst distributions shown in Figure 4D.
976 We considered three possible distributions: lognormal, Poisson, and negative-binomial, all with
977 unknown parameters. First, we simulated the viral burst size based on the assumed distribution.
978 For each simulated value of burst size x, we introduced measurement noise by assuming a log-
979 normal distribution. The mean of the log-normal distribution was determined by the bias function
980 $B(x)$, and the standard deviation was determined by $\sigma(x)$. Next, we computed a density function
981 using a kernel density estimation with a Gaussian kernel to represent the resulting distribution.
982 We then used this density function to calculate the log-likelihood of the observed data. We
983 estimated the parameters for each distribution and for each dataset (H3N2 or H1N1) by
984 maximizing the likelihood of observations reported in Fig. 4D.
985
986
987

References for supplementary information

1. Shu B, Wu KH, Emery S, Villanueva J, Johnson R, Guthrie E, et al. Design and performance of the CDC real-time reverse transcriptase PCR swine flu panel for detection of 2009 A (H1N1) pandemic influenza virus. J Clin Microbiol. 2011 Jul;49(7):2614–9.

2. Piorkowski G, Baronti C, de Lamballerie X, de Fabritus L, Bichaud L, Pastorino BA, et al. Development of generic Taqman PCR and RT-PCR assays for the detection of DNA and mRNA of β-actin-encoding sequences in a wide range of animal species. J Virol Methods. 2014 Jun;202:101–5.

3. Tao Y, Rotem A, Zhang H, Chang CB, Basu A, Kolawole AO, et al. Rapid, targeted and culture-free viral infectivity assay in drop-based microfluidics. Lab Chip. 2015 Sep;15(19):3934–40.

4. Mazutis L, Gilbert J, Ung WL, Weitz DA, Griffiths AD, Heyman JA. Single-cell analysis and sorting using droplet-based microfluidics. Nat Protoc. 2013 May;8(5):870–91.

5. Loveday EK, Zath GK, Bikos DA, Jay ZJ, Chang CB. Screening of Additive Formulations Enables Off-Chip Drop Reverse Transcription Quantitative Polymerase Chain Reaction of Single Influenza A Virus Genomes. Anal Chem. 2021 Mar;93(10):4365–73.

6. Gevertz JL, Dunn SM, Roth CM. Mathematical model of real-time PCR kinetics. Biotechnol Bioeng. 2005 Nov;92(3):346–55.

7. Lievens A, Van Aelst S, Van den Bulcke M, Goetghebeur E. Enhanced analysis of real-time PCR data by using a variable efficiency model: FPK-PCR. Nucleic Acids Res. 2012 Jan;40(2):e10.

8. Rutledge RG. Sigmoidal curve-fitting redefines quantitative real-time PCR with the prospective of developing automated high-throughput applications. Nucleic Acids Res. 2004 Dec;32(22):e178.

9. Tellinghuisen J, Spiess AN. qPCR data analysis: Better results through iconoclasm. Biomol Detect Quantif. 2019 Mar;17:100084.

10. Spiess AN, Feig C, Ritz C. Highly accurate sigmoidal fitting of real-time PCR data by introducing a parameter for asymmetry. BMC Bioinformatics. 2008 Apr;9:221.

11. Alvarez MJ, Vila-Ortiz GJ, Salibe MC, Podhajcer OL, Pitossi FJ. Model based analysis of real-time PCR data from DNA binding dye protocols. BMC Bioinformatics. 2007 Mar;8:85.

12. Loveday EK, Sanchez HS, Thomas MM, Chang CB. Single-Cell Infection of Influenza A Virus Using Drop-Based Microfluidics. Microbiol Spectr. 2022 Sep;e0099322.

13. Rutledge RG, Stewart D. A kinetic-based sigmoidal model for the polymerase chain reaction and its application to high-capacity absolute quantitative real-time PCR. BMC Biotechnol. 2008 May;8:47.

14. Ruijter JM, Pfaffl MW, Zhao S, Spiess AN, Boggy G, Blom J, et al. Evaluation of qPCR curve analysis methods for reliable biomarker discovery: bias, resolution, precision, and implications. Methods. 2013 Jan;59(1):32–46.

APPENDIX B

SUPPLEMENTAL INFORMATION FOR CHAPTER 5

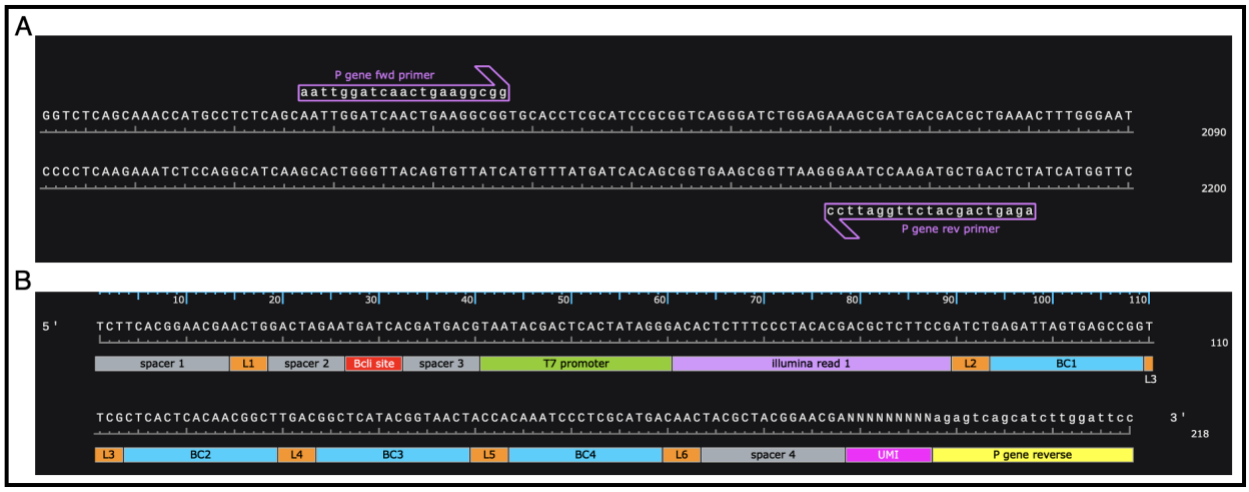


Figure 10MV P gene specific barcoding oligonucleotide primer (BOP). (A) The MV P gene contains attenuation markers within a 140 nt region of interest. Custom PCR primers have been validated to amplify this region. (B) For scRNA-seq of P gene mRNA, we construct hydrogel beads carrying barcoding oligonucleotide primers (BOP) that include the P gene reverse primer. BOPs also include sequences for BclI restriction enzymes, T7 promoter, Illumina read 1, a combination of four unique barcodes, and a read specific unique molecular identifier.

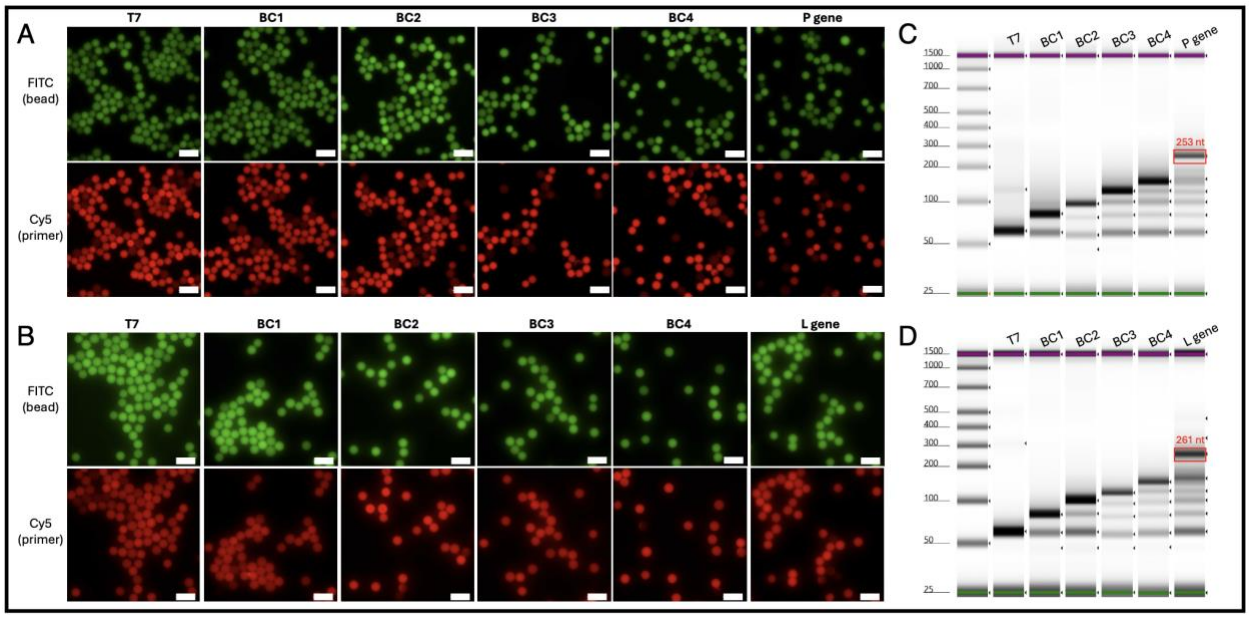


Figure 11. Control BHBs Synthesized with FITC-dextran and a Single Barcode Sequence. BHB synthesis protocols were validated with two GSP sequences, for MV P gene and MV L gene mRNA capture. Both types of. BHBs were made with FITC-dextran and a single barcode sequence. The quality of BHB synthesis was validated with (A) and (B) Fluorescence in situ hybridization (FISH) with Cy5 probes complementary to the T7 promoter site, barcode sequences 1 to 4, and either the AMV P gene mRNA primer site or L gene mRNA primer site. And (C) and (D) Capillary electrophoresis of ligation intermediate samples, used to measure the length of each ligation product.

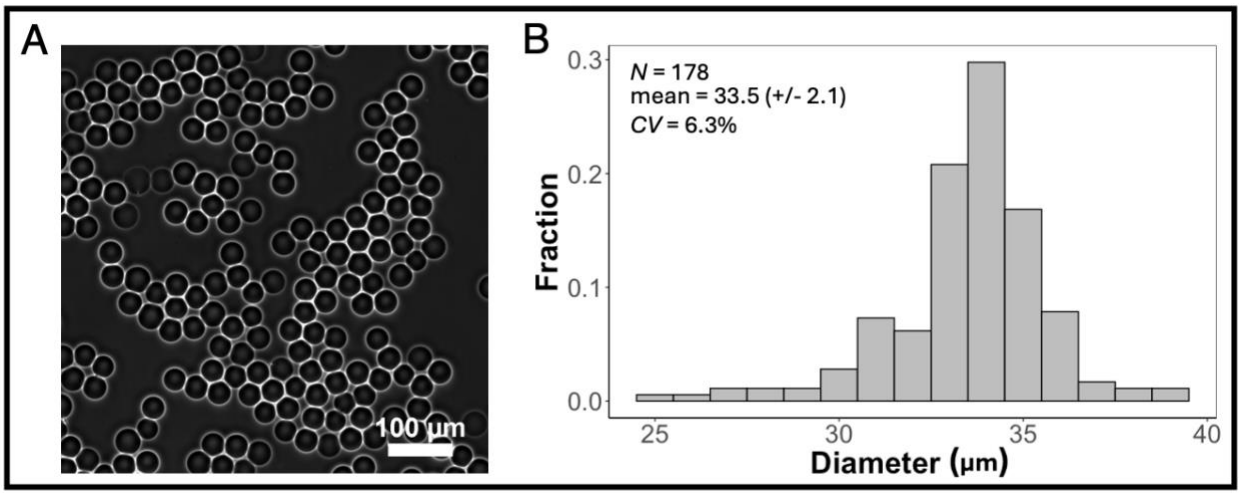
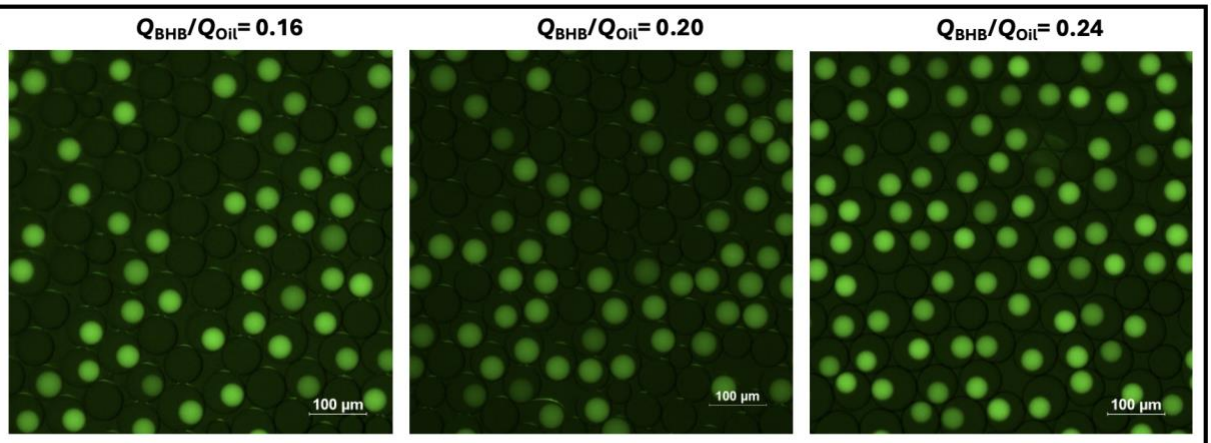


Figure 12. Hydrogel Bead (HB) Size. (A) Hydrogel beads (HBs) used in device operation. (B) The average diameter of HBs (N = 178 beads) was 33.5 µm (+/- 2.1) with a CV= 6.3% , measured using Nikon image analysis software.

A



B

condition	total drops	0	1	2	3
Q_{BHB}/Q_{Oil}= 0.16	106	0.452830189	0.547169811	0	0
Q_{BHB}/Q_{Oil}= 0.20	130	0.446153846	0.553846154	0	0
Q_{BHB}/Q_{Oil}= 0.24	105	0.142857143	0.857142857	0	0

Figure 13. BHB Loading by Flow Rate Fraction. (A) Control BHBs were loaded into the co-encapsulation device at three flow rate fractions (QBHB/Qoil), including QBHB/Qoil = 0.16, QBHB/Qoil=0.20, and QBHB/Qoil=0.24 then imaged off chip under fluorescence. (B) The number of BHBs per drop were counted from each image. The flow rate fraction QBHB/Qoil=0.24 gave the greatest proportion of single BHBs per drop and was used for the final device operation protocol.

A

condition	total drops	0	1	2	3
2x10⁶ cells/mL	152	0.486842105	0.368421053	0.131578947	0.013157895
1x10⁶ cells/mL	234	0.598290598	0.320512821	0.05982906	0.017094017

B

cells/drop	Probability
0	0.6065307
1	0.3032653
2	0.0758163
3	0.0126361
4	0.0015795
5	0.000158
6	1.316E-05
7	9.402E-07
8	5.876E-08
9	3.265E-09
10	1.632E-10
11	7.419E-12
12	3.091E-13
13	1.189E-14
14	4.246E-16
15	1.415E-17
16	4.423E-19
17	1.301E-20
18	3.614E-22
19	9.51E-24
20	2.378E-25

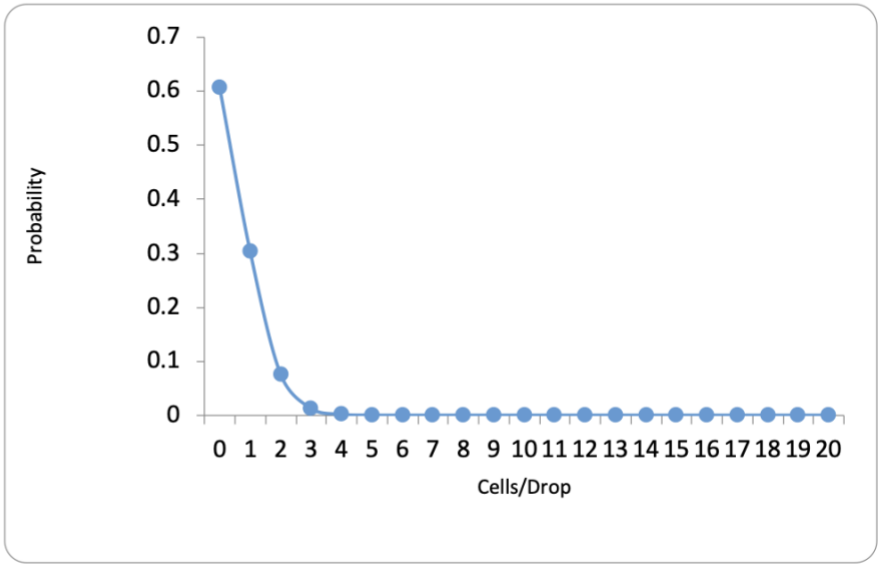


Figure 14. Single-cell Loading by Suspension Concentration. (A) MV infected Vero cells were loaded into the co-encapsulation device at two cell suspension concentrations, 2x106 cells/mL and 1x106 cells/mL then imaged off chip to count the number of cells per drop. The loading concentration of 1x106 cells/mL gave the greatest proportion of single-cells per drop and was used for the final device operation protocol. This loading distribution followed the (B) Poisson distribution for loading 0.5 cells per drop, using a 1x106 cells/mL suspension encapsulated into 60μm droplets.

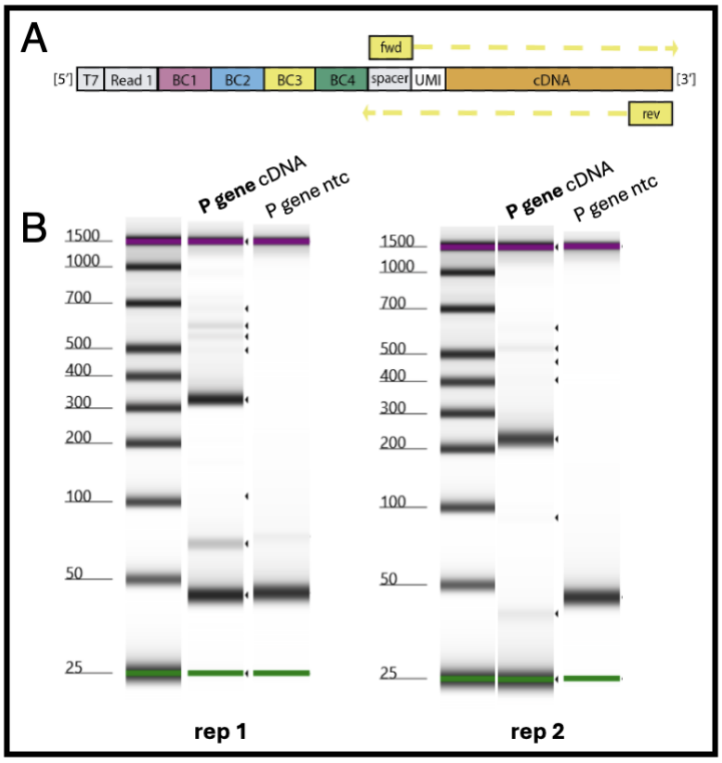


Figure 15. Barcoded cDNA (BC-cDNA) average nucleotide length. (A) BC-cDNA was detected using a custom qPCR assay with forward primers targeting a spacer region between the barcode 4 and UMI sequences, and reverse primers targeting the 3' end of cDNA. The qPCR products for two scRNA-seq independent replicates are shown in the main text (Fig 3F) (B) The nucleotide length of BC-cDNA was measured with capillary electrophoresis and measured between 200 and 300 nt on average.

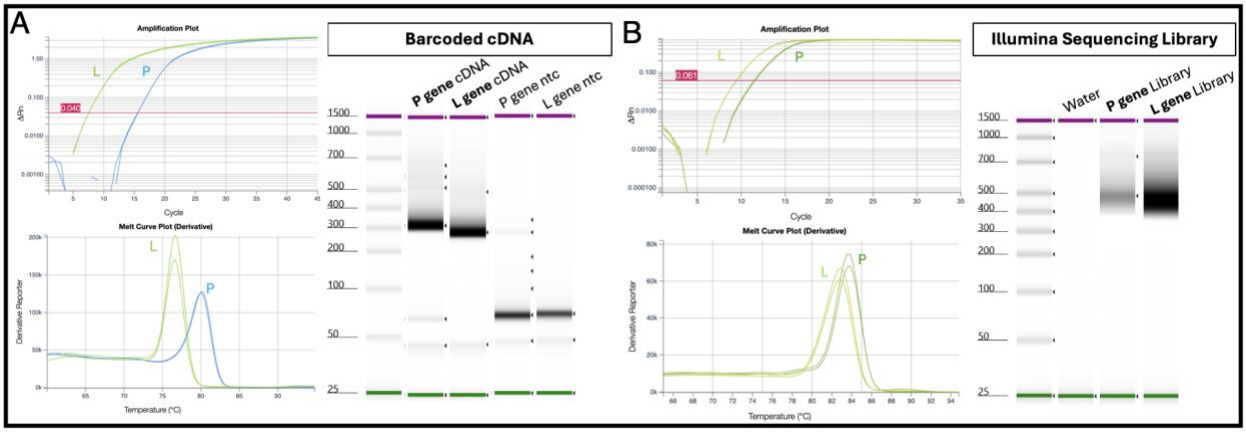


Figure 16. Protocol validation with single-barcode BHBs and synthetic RNA template. Control experiments were conducted using BHBs carrying a single barcode sequence across all beads. (A) Single barcode BHBs were tested for cDNA synthesis and barcoding function using synthetic RNA templates with sequences for the MV L and P gene mRNA. Barcoded cDNA products were detected with a custom qPCR assay, and nucleotide length of each qPCR product was measured by capillary electrophoresis using the Tapestation D1000 kit. (B) The product of each Illumina sequencing library was detected by qPCR, using the KAPPA Illumina Library Quantification Kit. The length of each Illumina sequencing library was measured by capillary electrophoresis, using the Tapestation HSD1000 kit.

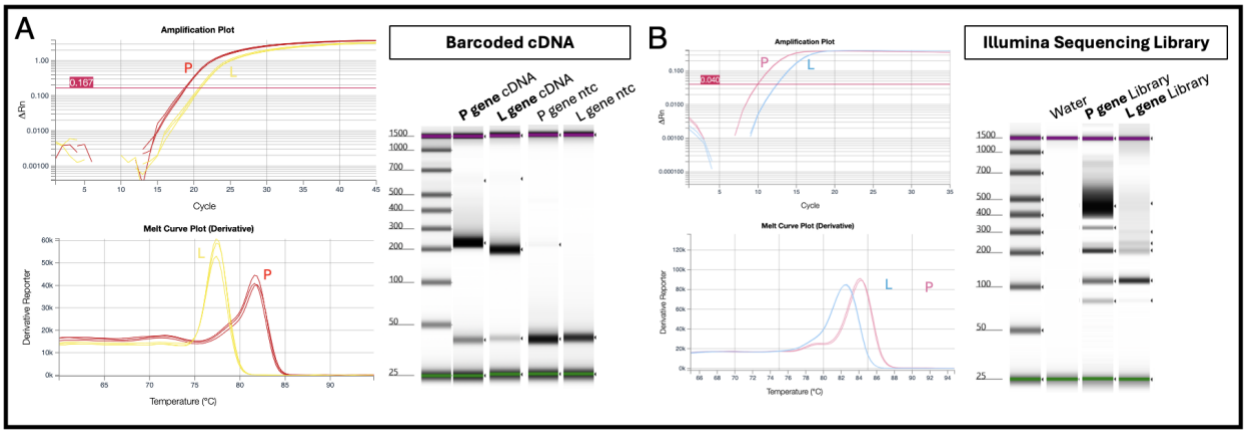


Figure 17. Protocol validation with single-barcode BHBs and MV infected cells in bulk. Control experiments were conducted using BHBs carrying a single barcode sequence across all beads. (A) Single barcode BHBs were tested for cDNA synthesis and barcoding function of P gene and L gene mRNA from MV infected cells in bulk. Barcoded cDNA products were detected with a custom qPCR assay, and nucleotide length of each qPCR product was measured by capillary electrophoresis using the Tapestation D1000 kit. (B) The product of each Illumina sequencing library was detected by qPCR, using the KAPPA Illumina Library Quantification Kit. The length of each Illumina sequencing library was measured by capillary electrophoresis, using the Tapestation HSD1000 kit.

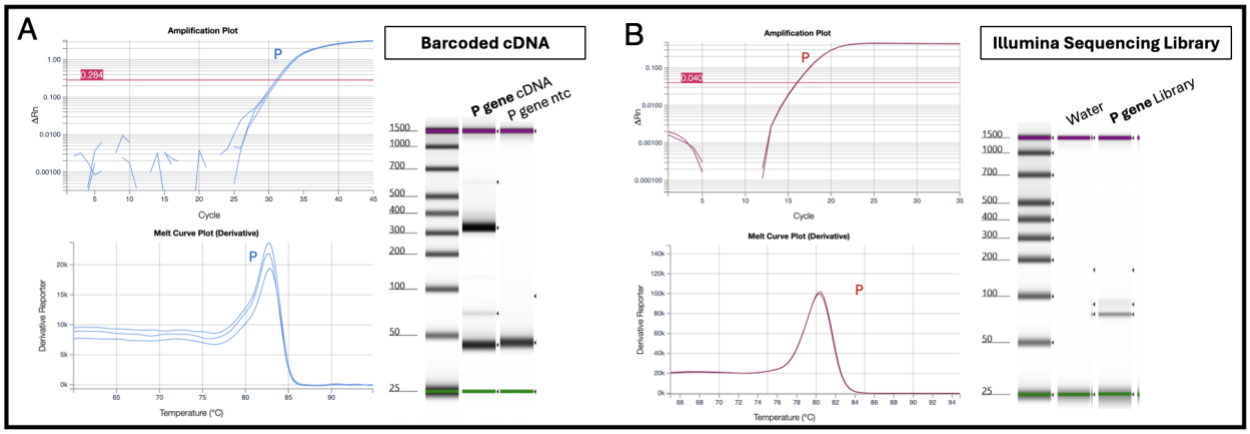


Figure 18. Protocol validation with single-barcode BHBs and MV infected cells in drops. Control experiments were conducted using BHBs carrying a single barcode sequence across all beads. (A) Single barcode BHBs were tested for cDNA synthesis and barcoding function of P gene mRNA from MV infected cells in drops. Barcoded cDNA products were detected with a custom qPCR assay, and nucleotide length of each qPCR product was measured by capillary electrophoresis using the Tapestation D1000 kit. (B) The product of each Illumina sequencing library was detected by qPCR, using the KAPPA Illumina Library Quantification Kit. The length of each Illumina sequencing library was measured by capillary electrophoresis, using the Tapestation HSD1000 kit.

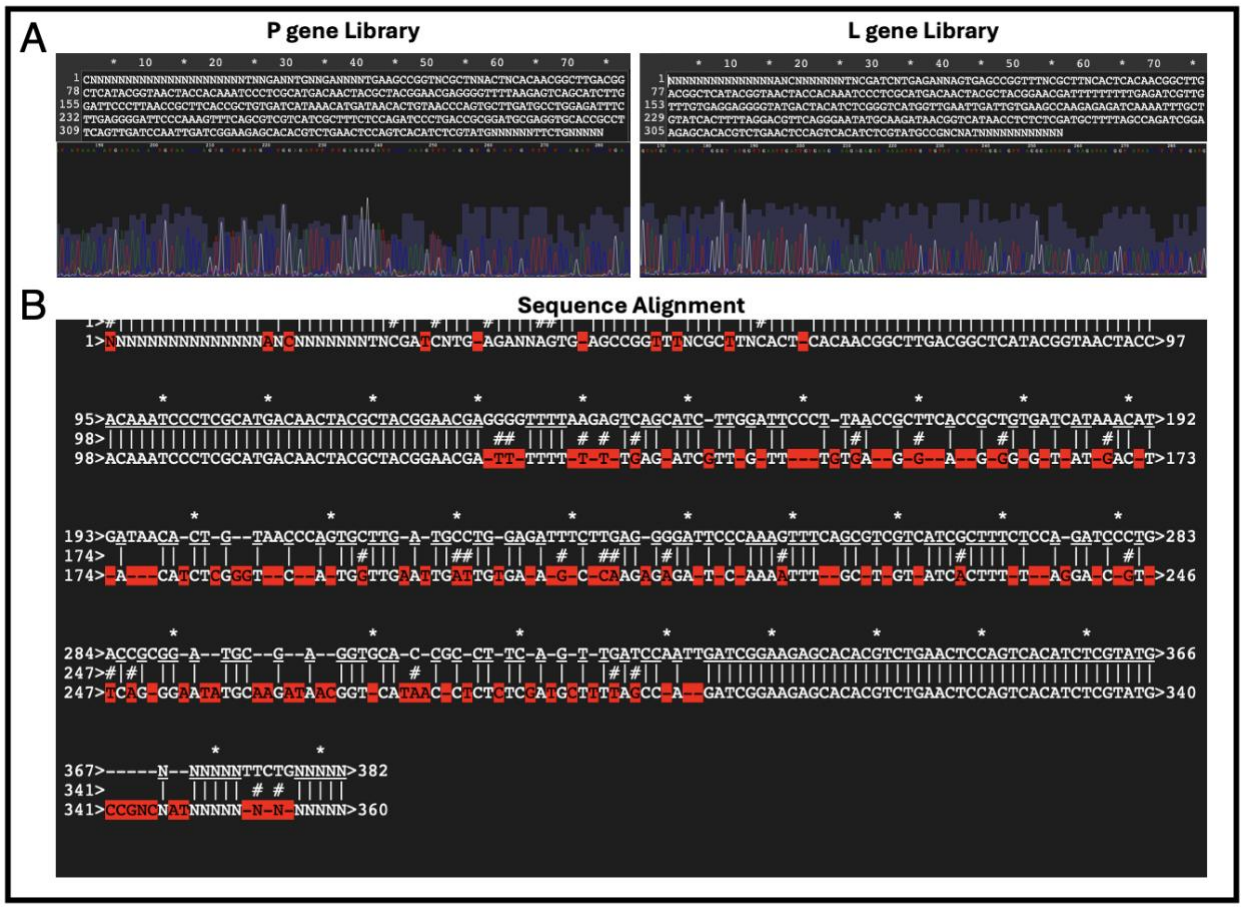


Figure 19. Sanger Sequencing of Library from Single Barcode BHBs + Synthetic RNA Gene Block Template in Bulk. (A) Libraries were sequenced using the Sanger method, to verify the presence of known barcode and mRNA sequences. (B) Sequence alignment of both libraries shows shared barcode and Illumina sequences on the 5' and 3', and different mRNA inserts (P or L gene) in the middle of the sequence.

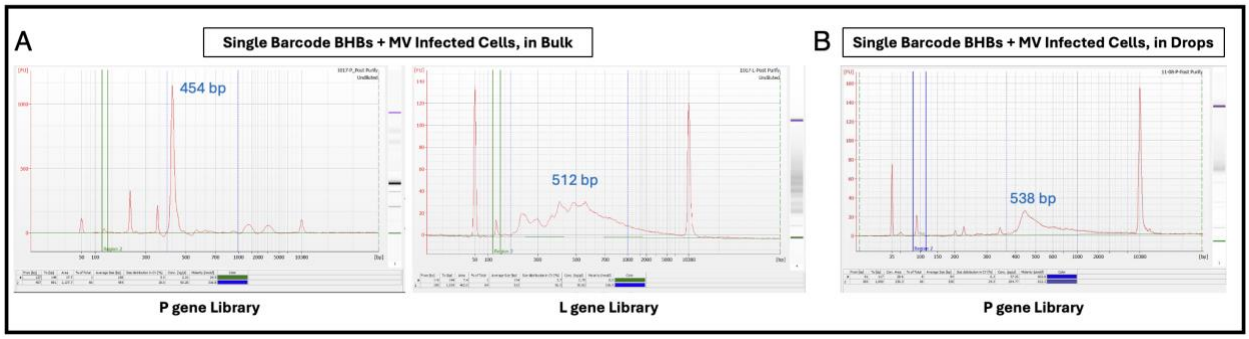


Figure 20. Libraries from Single Barcode BHBs Controls. The average length of each Illumina sequencing library was also measured using the Bioanlyzer HSD1000 kit. (A) Protocol validation with single-barcode BHBs and MV infected cells in bulk showed libraries averaging 454 nt (P gene) and 512 nt (L gene) in length (B) Protocol validation with single-barcode BHBs and MV infected cells in drops showed libraries averaging 538 nt in length.

Table 1. Barcoding Oligonucleotide Primer (BOP) Sequences.

Primer Name and Type	Primer Features	Strand Name	Sequence (5' to 3')
Primer 1 dsDNA	Acrydite-modified for hydrogel integration.	Primer 1 fwd	/Acryd/TCTTCACG GAACGA
		Primer 1 rev	/Phos/CAGT TCGTTCCGTGAAG A
Primer 2 dsDNA	BclI restriction enzyme site, T7 promoter, Illumina read 1.	Primer 2 fwd	/Phos/ACTGGACTA GAATGATCACGAT GACGTAATACGA CTCACTATAGGGA CACTCTTTCCCTA CACGACGCTCTTC CG
		Primer 2 rev	AGATCGGAAGAG CGTCGTGTAGGGA AAGAGTGTCCCTA TAGTGAGTCGTAT TACGTCATCGTGA TCATTCTAGTC
Primer 3 dsDNA	Barcode 1 for control beads	Primer 3 fwd	/Phos/ATCTGAGAT TAGTGAGCCGG
		Primer 3 rev	CGAACCGGCTCAC TAATCTC
Primer 4 dsDNA	Barcode 2 for control beads	Primer 4 fwd	/Phos/TTCGCTCAC TCACAACGGCT
		Primer 4 rev	GTCAAGCCGTTGT GAGTGAG
Primer 5 dsDNA	Barcode 3 for control beads	Primer 5 fwd	/Phos/TGAC GGCTCATACGGTA ACT
		Primer 5 rev	TGGTAGTTACCGT ATGAGCC

Primer 6 dsDNA	Barcode 4 for control beads	Primer 6 fwd	/Phos/ACCACAAATCCCTCGCATGA
		Primer 6 rev	GTTGTCATGCGAGGGATTTG
Primer 7 ssDNA	Unique molecular identifier, P gene primer for genome.	Primer 7 P gene fwd	/Phos/CAACTACGCTACGGAACGANNNNNNNNNAATGGATCAACTGAAGGCGG
	Unique molecular identifier, P gene primer for mRNA.	Primer 7 P gene rev	/Phos/CAACTACGCTACGGAACGANNNNNNNNNAGAGTCAGCATCTTGGATTCC
	Unique molecular identifier, L gene primer for genome.	Primer 7 L gene fwd	/Phos/CAACTACGCTACGGAACGANNNNNNNNNTGGCTAAAAGCATCGAGAGAG
	Unique molecular identifier, L gene primer for mRNA.	Primer 7 L gene rev	/Phos/CAACTACGCTACGGAACGANNNNNNNNNGAGGTCGTTGTTTGTGAGGAG
Primer 8 ssDNA	BclI complement, T7 complement.	Primer 8	CCCTATAGTGAGTCGTATTACGTCATCGTGATCATTCTAGTC

The sequences for Primers 3 to 6 (Barcodes 1 to 4) can be found in an attached file. There are 96 sequences provided for each barcode type, which gives 96^4 (8.5×10^7) unique barcodes after the split and pool ligation procedure. This is roughly equivalent to the number of BHBs (4×10^7) for which this protocol is written. Each primer was ordered in two separate 96 well plates, one for the forward strand and one for the reverse strand. To anneal the two strands before use in split and pool ligation, 15 µL of 500 µM forward primer was combined with 15 µL of reverse primer. Ensuring that well positions (A1 forward + A1 reverse) were constant. Annealing was completed under the following conditions: 98°C for 15 seconds; Decrease temperature by 1°C every 20 seconds until 15°C (20s at 97°C > 20s 96°C > 20s 95°C…. > 20s 15°C). The final concentration was 250 µM for annealed dsDNA sequences.

Table 2. BOP FISH Probe Sequences.

FISH Probe Target Region	Sequence (5’ to 3’)
Primer 2 – T7 promoter	/Cy5/CCCTATAGTGAGTCGTATTA
Primer 2 – Illumina read 1	/Cy5/CGGAAGAGCGTCGTGTAGGGAAAGAGTGT
Primer 3 – Barcode 1	/Cy5/CCGGCTCACTAATCTC
Primer 4 – Barcode 2	/Cy5/AGCCGTTGTGAGTGAG
Primer 5 – Barcode 3	/Cy5/AGTTACCGTATGAGCC
Primer 6 – Barcode 4	/Cy5/TCATGCGAGGGATTTG
Primer 7 – P gene forward	/Cy5/CCGCCTTCAGTTGATCCAATT
Primer 7 – P gene reverse	/Cy5/GGAATCCAAGATGCTGACTCT
Primer 7 – L gene forward	/Cy5/CTCTCTCGATGCTTTTAGCCA
Primer 7 – L gene reverse	/Cy5/CTCCTCACAAACAACGACCTC

Table 3. PCR Primers for Detection of MV Barcoded cDNA.

Primer	Sequence (5’ to 3’)
BC/UMI forward	CAACTACGCTACGGAACGA
P gene forward	AATTGGATCAACTGAAGGCGG
P gene reverse	AGAGTCAGCATCTTGGATTCC
L gene forward	TGGCTAAAAGCATCGAGAGAG
L gene reverse	GAGGTCGTTGTTTGTGAGGAG

Table 4. Sequencing Library Preparation Primers.

Primer Name and Type	Primer Features	Strand Name	Sequence (5' to 3')
Primer 9 ssDNA	Illumina read 2, P gene primer for genome. *Use with Primer 7 P gene rev.*	Primer 9 P gene fwd	GTGACTGGAGTTCAGACGTGT GCTCTTCCGATCAATTGGATCA ACTGAAGGCGG
	Illumina read 2, P gene primer for mRNA. *Use with Primer 7 P gene fwd.*	Primer 9 P gene rev	GTGACTGGAGTTCAGACGTGT GCTCTTCCGATCAGAGTCAGC ATCTTGGATTCC
	Illumina read 2, L gene primer for genome. *Use with Primer 7 L gene rev.*	Primer 9 L gene fwd	GTGACTGGAGTTCAGACGTGT GCTCTTCCGATCTGGCTAAAA GCATCGAGAGAG
	Illumina read 2, L gene primer for mRNA. *Use with Primer 7 L gene fwd.*	Primer 9 L gene rev	GTGACTGGAGTTCAGACGTGT GCTCTTCCGATCGAGGTCGTTG TTTGTGAGGAG
Primer 10 ssDNA	Illumina P5 adaptor	Primer 10	AATGATACGGCGACCACCGAG ATCTACACTCTTTCCCTACACG ACG
Primer 11 ssDNA	Illumina P7 adaptor	Primer 11	CAAGCAGAAGACGGCATACGA GATGTGACTGGAGTTCAGACG TGTG

Table 5. Sequencing Library Quantification Results using Qubit Assay.

GSP	BHBs	RNA Target	Method	Replicate	Concentration (ng/ul)
P gene mRNA	Single Barcode	MV infected Cells	Bulk	Rep 1	6.6
L gene mRNA	Single Barcode	MV infected Cells	Bulk	Rep 1	0.549
P gene mRNA	Single Barcode	MV infected Cells	Drops	Rep 1	0.0133
P gene mRNA	Multi Barcode	MV infected Cells	Drops	Rep 1	0.781
P gene mRNA	Multi Barcode	MV infected Cells	Drops	Rep 2	0.0369
P gene mRNA	Multi Barcode	MV infected Cells	Drops	Rep 3	0.0214

Table 6. Synthetic RNA Gene Block Sequences.

Gene Block (ssDNA)	Sequence (5’ to 3’)
measles virus P gene	cagcatggtcagaaatatcagacaacccaggacaggaccgagccacctgcaag gaagaggaggcaggcagttcgggtctcagcaaaccatgcctctcagcaattggat caactgaaggcggtgcacctcgcatccgcggtcagggatctggagaaagcgatg acgacgctgaaactttgggaatcccctcaagaaatctccaggcatcaagcactgg gttacagtgttatcatgtttatgatcacagcggtgaagcggttaagggaatccaagat gctgactctatcatggttcaatcaggccttgatggtgatagcaccctctcaggagga gacgatgaatctgaaaacagcgatgtggatattggcgaacctgataccgagggat atgctatcactga
measles virus L gene	aaggaatatattatgatgggctacttgtgtcccaatcactcaagagcatcgcaagat gtgtattctggtcagagactatagttgatgaaacaagggcagcatgcagtaatattg ctacaacaatggctaaaagcatcgagagaggttatgaccgttatcttgcatattccct gaacgtcctaaaagtgatacagcaaattttgatctctcttggcttcacaatcaattcaa ccatgacccgagatgtagtcatacccctcctcacaaacaacgatctcttaataagga tggcactgttgcccgctcctattgggggatgaattatctgaatatgagcaggctgtt tgtcagaaacatcggtgatccagtaacatcatcaattgctgatctcaagagaat

Table 7. 1M NaCl Solution.

Reagent	Starting Concentration	Volume Added	Final Concentration
NaCl	5 M	40 ml	1 M
Nuclease-free water	na	160 ml	na

(200 ml volume). Prepare in 500 ml sterile plastic media bottle. Filter the solution through a 0.2-µm membrane. Wrap bottle in parafilm and store at room temperature for up to 6 months.

Table 8. 10% Triton Solution.

Reagent	Starting Concentration	Volume Added	Final Concentration
Triton X-100	100%	5 ml	10%
Nuclease-free water	na	45 ml	na

(50 ml volume). Prepare in 50 ml nuclease-free tube. Rotate for 1-2 hours to mix. Wrap tube in parafilm and store at room temperature for up to 6 months.

Table 9. 1X TBSET Buffer Solution.

Reagent	Starting Concentration	Volume Added	Final Concentration
Tris HCl pH 8.0	1 M	5 ml	10 mM
NaCl	1 M	68.5 ml	137 mM
KCl	2 M	675 ul	2.7 mM
EDTA pH 8.0	0.5 M	10 ml	10 mM
Triton	10%	5 ml	0.10%
Nuclease-free water	na	410.825 ml	na

(500 ml volume). Prepare in 500 ml sterile plastic media bottle. Filter the solution through a 0.2-µm membrane. Wrap bottle in parafilm and store at room temperature for up to 6 months.

Table 10. 4X AB Crosslinking Agent Solution.

Reagent	Starting Concentration	Volume Added	Final Concentration
acrylamide/bis-acrylamide	40%	3.6 ml	14.40%
acrylamide	40%	2.58 ml	10.32%
Nuclease-free water	na	3.82 ml	na

(10 ml volume). *Acrylamide and bis-acrylamide are potent neurotoxins. Use in well-ventilated areas and wear PPE.* Prepare in 50 ml nuclease-free tube. Filter the solution through a 0.2-µm membrane. Transfer to a 15 ml nuclease-free tube. Top with N2 gas. Wrap tube in parafilm and store at 4 °C for up to 6 months.

Table 11. 10% APS Solution.

Reagent	Starting Concentration	Volume Added	Final Concentration
APS	na	1 g	na
Nuclease-free water	na	10 ml	na

(10 ml volume). *APS is an oxidizing agent that may ignite combustible materials. Is harmful if swallowed. Wear PPE.* Prepare in 50 ml nuclease-free tube. Filter the solution through a 0.2-µm membrane. Aliquot 500 µl into 1.5 ml nuclease-free tubes. Store at −20 °C for up to 6 months.

Table 12. 20% PFO Solution.

Reagent	Starting Concentration	Volume Added	Final Concentration
PFO	100%	2 ml	10%
HFE7500	na	8 ml	na

(10 ml volume). Prepare in 15 ml nuclease-free tube. Filter the solution through a 0.2-µm membrane, for oil. Wrap tube in parafilm and store at room temperature for up to 6 months.

Table 13. 1% Span 80 Solution.

Reagent	Starting Concentration	Volume Added	Final Concentration
Span 80	100%	200 ul	10%
Hexane	na	19.8 ml	na

(20 ml volume). *Hexane is highly flammable and toxic. Prepare and store solution in glass bottle in the fume hood* Prepare in 100 ml glass bottle. Mix on stir plate for 1-2 hours. Store at room temperature in fume hood for up to 6 months.

Table 14. 5% (w/w) RAN Solution.

Reagent	Starting Concentration	Volume Added	Final Concentration
RAN	na	0.815 g	5 wt%
HFE7500	na	16.3 g (10 ml)	na

(10 ml volume). Filter HFE7500 through a 0.2-µm membrane, for oil. Calculate the mass needed of HFE7500 and RAN. Resuspend neat RAN stock in filtered HFE7500. Add a 15ml nuclease-free tube to the balance, and tare. Measure out 0.815 g RAN. Remove from balance and add 10 ml HFE7500. Rotate overnight to mix. Wrap in parafilm and store at 4°C.
Calculate HFE7500 mass in 10ml solution. HFE7500 density is 1.63 g /mL

$$10\ ml\ \times\ \frac{1.63\ g}{1\ ml} \rightarrow \mathbf{16.3\ g\ HFE7500}$$

Calculate RAN mass in a 5% (w/w) solution.

$$C_1W_{1=}C_2W_2$$
$$(100\%)(W_1) = (5\%)(16.3\ g)\ \rightarrow \mathbf{0.815\ g\ RAN}$$

Table 15. 3% RAN Solution.

Reagent	Starting Concentration	Volume Added	Final Concentration
RAN	5%	3 ml	3%
HFE7500	na	2 ml	Na

(5ml volume). Prepare in 15 ml nuclease-free tube. Filter the solution through a 0.2-µm membrane, for oil. Store at 4C.

Table 16. HB Synthesis Aqueous Phase Solution.

Reagent	Starting Concentration	Volume Added	Final Concentration
TBSET	1x	50 ul	0.1x
AB	4x	125 ul	1x
APS	10%	15 ul	0.30%
Primer 1	645uM	310 ul	400uM

(500ul volume).

Table 17. HB Synthesis Oil Phase Solution.

Reagent	Starting Concentration	Volume Added	Final Concentration
RAN in HFE7500	2%	2.5 ml	2%
TEMED	100%	10 ul	0.4%

(2.5ml volume).

Table 18. Primer 2 Ligation Mixture Solution.

Reagent	Starting Concentration	Volume Added	Final Concentration
Primer 2	500uM	800 ul	130uM
NEB Quick Ligation Buffer	5x	600 ul	1x
NEB T7 DNA Ligase	3000 U/ul	30 ul	30U/ul
Nuclease-free water	na	1570 ul	na

(3ml volume).

Table 19. Primer 3 to 6 Ligation Mixture Solution.

Reagent	Starting Concentration	Volume Added	Final Concentration
NEB Quick Ligation Buffer	5x	600 ul	1.4x
NEB T7 DNA Ligase	3000 U/ul	30 ul	30U/ul

(630ul volume).

Table 20. Primer 7 Ligation Mixture Solution.

Reagent	Starting Concentration	Volume Added	Final Concentration
Primer 7	500 uM	420 ul	70 uM
NEB Quick Ligation Buffer	5x	600 ul	1x
NEB T7 DNA Ligase	3000 U/ul	30 ul	30U/ul
Nuclease-free water	na	1950 ul	na

(3ml volume).

Table 21. 10% Tween 20 Solution.

Reagent	Starting Concentration	Volume Added	Final Concentration
Tween 20	100%	2 ml	10%
Nuclease-free water	na	18 ml	na

(20 ml volume). Prepare in a 50 ml nuclease-free tube. Filter the solution through a 0.2-µm membrane. Wrap in parafilm and store at room temperature for up to 6 months.

Table 22. Hydrogel Bead Wash (HBW) Buffer Solution.

Reagent	Starting Concentration	Volume Added	Final Concentration
Tris-HCl pH 8.0	1 M	5 ml	10 mM
EDTA	0.5 M	100 ul	1 mM
Tween 20	10 % (vol/vol)	5 ml	0.1%
Nuclease-free water	na	490 ml	na

(500 ml volume). Prepare in a 500 ml sterile plastic media bottle. Filter the solution through a 0.2-µm membrane. Wrap in parafilm and store at room temperature for up to 6 months.

Table 23. Denaturation Buffer Solution.

Reagent	Starting Concentration	Volume Added	Final Concentration
NaOH	10 M	750 ul	150 mM
Brij 30	30% (w/w)	833 ul	0.5%
Nuclease-free water	na	48.417 ml	na

(50 ml volume). *NaOH is highly corrosive. Wear PPE to protect skin.* Prepare on day of experiment - in a 50 ml nuclease-free tube. Brij 30 will gel at room temperature, incubate in 70C water bath for 30 min to convert to liquid phase. Filter the solution through a 0.2-µm membrane. Do not store, discard unused volume.

Table 24. Neutralization Buffer Solution.

Reagent	Starting Concentration	Volume Added	Final Concentration
Tris-HCl pH 8.0	1 M	50 ml	100 mM
NaCl	1 M	50 ml	100 mM
EDTA	0.5 N	10 ml	10 mM
Tween 20	10 % (vol/vol)	5 ml	0.1%
Nuclease-free water	na	385 ml	na

(500 ml volume). Prepare in a 500 ml sterile plastic media bottle. Filter the solution through a 0.2-µm membrane. Wrap in parafilm and store at room temperature for up to 6 months.

Table 25. Hybridization Buffer Solution.

Reagent	Starting Concentration	Volume Added	Final Concentration
Tris-HCl pH 8	1 M	5 ml	10 mM
KCL	2 M	82.5 ml	330.3 mM
EDTA	0.5 M	100 ul	0.1 mM
Tween 20	10% (vol/vol)	5 ml	0.1%
Nuclease-free water	na	407 ml	na

(500 ml volume). Prepare in 500 ml sterile plastic media bottle. Filter the solution through a 0.2-µm membrane. Wrap in parafilm and store at room temperature for up to 6 months.

Table 26. BHB Pre-Encapsulation Wash Solution.

Reagent	Starting Concentration	Volume Added	Final Concentration
Triton X100	10%	3 ml	0.60%
Tris-HCL pH 7.4	1 M	7.5 ml	150 mM
Nuclease-free water	na	39.5 ml	na

(50 ml volume). Prepare on day of experiment in a 50 ml nuclease-free tube. Filter through 0.2um membrane, for aqueous phases, prior to use. Do not store, discard unused volume.

Table 27. BHB Activation Wash Solution.

Reagent	Starting Concentration	Volume Added	Final Concentration
Triton X100	10%	10 ul	0.60%
Tris-HCL pH 7.4	1 M	30 ul	150 mM
DTT	0.1M	30 ul	15mM
FSS Buffer	5x	100 ul	2.5x
Nuclease-free water	na	28 ul	na

(200ul volume). Prepare on day of experiment - in a 1.5ml nuclease-free tube. Do not store, discard unused volume.

Table 28. Enzyme Mix Solution.

Reagent	Starting Concentration	Volume Added	Final Concentration
FSS buffer	5x	125 ul	1.25x
DTT	0.1M	25 ul	5mM
dNTP	10mM	75 ul	1.5mM
Triton X100	10%	45 ul	0.90%
Superscript III RT	200 U/ul	62.5 ul	25 U/ul
BcII	10,000 U/ml	53.5 ul	1.07 U/ul
SUPERaseln	20 U/ul	62.5 ul	2.5 U/ul
Nuclease-free water	na	51.5 ul	na

(500ul volume).

Table 29. 80% EtOH Solution.

Reagent	Starting Concentration	Volume Added	Final Concentration
EtOH	100%	32 ml	80%
Nuclease-free Water	na	8 ml	na

(40 ml volume). Prepare in 50ml nuclease-free tube on day of experiment. Caution: 100% EtOH is extremely flammable. Do not store, discard after use.

Table 30. gBlock Resuspension Solution.

Reagent	Starting Concentration	Volume Added	Final Concentration
L or P gene G Block	3ug of a 400bp sequence	300 ul nuclease free water	10ng /ul

Order 400bp fragments that include L and P gene forward and reverse primer regions. IDT prepares ssDNA megamers normalized to a 3ug yield. 3ug = 24 pmol. 24pmol of 400 nt ssDNA = 10^{13} copies ($1.46x10^{13}$). Resuspend lyophilized pellet at 10ng/ul concentration. 3ug = 3000ng. Calculate, (10 ng / 1 ul) = (3000 ng / ? ul) → 300ul. Adding 300ul of nuclease-free water to a 3000ng lyophilized pellet gives a final resuspension concentration of 10ng/ul. 10ng of 400 nt ssDNA = $4.867x10^{10}$ copies. So 10ng/ul = ~10^{10} copies/ul of gBlock. Aliquot into nuclease-free tubes. Store at -20C.

Table 31. DTT 10^{-1} Dilution Solution.

Reagent	Starting Concentration	Volume Added	Final Concentration
DTT stock	0.1 M (100mM)	10 ul	10mM
Nuclease-free water	na	90 ul	na

(100ul volume).

Table 32. DTT 10^{-2} Dilution Solution.

Reagent	Starting Concentration	Volume Added	Final Concentration
DTT stock	10mM	10 ul	1mM
Nuclease-free water	na	90 ul	na

(100ul volume).

Table 33. DTT 10^{-3} Dilution Solution.

Reagent	Starting Concentration	Volume Added	Final Concentration
DTT stock	1mM	10 ul	0.1mM
Nuclease-free water	na	90 ul	na

(100ul volume). Perform a 10-fold dilution series (100mM→10mM→1mM→0.1mM) to reach the final working concentration.

Table 34. Bulk cDNA Synthesis and Barcoding Assay.

Reagent	Starting Concentration	Volume Added (single reaction)	Volume Added (master mix)	Final Concentration
Optiprep	60%	6.4 ul	371.2 ul	19.2%
FSS Buffer	5X	4 ul	232 ul	1x
Triton X100	10%	1 ul	58 ul	0.5%
DTT	0.1 mM	1 ul	58 ul	5uM
dNTP	10mM	1 ul	58 ul	0.5mM
SUPERaseln nuclease inhibitor	20 U/ul	1 ul	58 ul	1 U/ul
Bc1I restriction endonuclease	10,000 U/ml	1 ul	58 ul	10 U/ml
Super Script III Reverse Transcriptase Enzyme	200 U/ul	1 ul	58 ul	10 U/ul
BHBs	na	1 ul	na	na
RNA	10^9 copies / ul	2.6 ul	na	na

(20ul single reactions; 48 reaction master mix).

Table 35. cDNA Enzymatic Cleanup Assay.

Reagent	Starting Concentration	Volume Added	Final Concentration
ExoI	20 U/ul	10 ul	6 U/ul
R3.1 Buffer	10x	30 ul	1x
cDNA sample	na	60 ul	na

(100ul reactions).

Table 36. Power SYBR qPCR Assay.

Reagent	Starting Concentration	Volume Added	Final Concentration
SYBR Green PCR Master Mix	2x	10 ul	1x
Fwd primer (UMI/BC seq)	10uM	0.5 ul	250nM
Rev primer (GSP fwd seq)	10uM	0.5 ul	250nM
Barcoded cDNA	na	2 ul	na
Nuclease-free water	na	7 ul	na

(20ul reactions).

Table 37. Second Strand Synthesis Assay.

Reagent	Starting Concentration	Volume Added	Final Concentration
cDNA	na	15 ul	na
SSS Reaction Buffer	na	2 ul	na
SSS Enzyme Mix	na	1 ul	na
Nuclease-free water	na	2 ul	na

(20 ul reactions).

Table 38. In Vitro Transcription – NTP/ Buffer Master Mix.

Reagent	Starting Concentration	Volume Added	Final Concentration
T7 reaction buffer	10X	100 ul	2X
ATP	100mM	100 ul	20 mM
GTP	100mM	100 ul	20 mM
UTP	100mM	100 ul	20 mM
CTP	100mM	100 ul	20 mM

(500ul volume).

Table 39. In Vitro Transcription Assay.

Reagent	Starting Concentration	Volume Added	Final Concentration
SSS dsDNA	na	8 ul	na
NTP/ buffer master mix	na	10 ul	na
T7 RNA Polymerase	na	2 ul	na

(20 ul reactions).

Table 40. Secondary Reverse Transcription Assay.

Reagent	Starting Concentration	Volume Added	Final Concentration
FSS Buffer	5x	4ul	1x
SSIII RT Enzyme	200 U/ul	1ul	10 U/ul
dNTP	10mM	1ul	0.5mM
SUPERaseln	20 U/ul	1 ul	1 U/ul
Primer 9	100uM	1ul	5uM
IVT RNA product	10 pg to 500 ng	12 ul	(500nM ideally)

(20ul reactions).

Table 41. PCR for Library Amplification and Adaptor Ligation Assay.

Reagent	Starting Concentration	Volume Added	Final Concentration
GC Buffer	5x	10 ul	1x
dNTP	10mM	1.5 ul	0.3mM
P5-Rd1 (primer 10)	10uM	2.5 ul	0.5uM
P7-id-Rd2 (primer 11)	10uM	2.5 ul	0.5uM
KAPA Polymerase	1U/ul	0.5 ul	0.5U
Nuclease-free water	na	28 ul	na
cDNA from RT2	na	5 ul	na

(50ul reactions).

Table 42. Primer/Dye Master Mix for Illumina Library qPCR Assay.

Reagent	Starting Concentration	Volume Added	Final Concentration
KAPA SYBR FAST qPCR Master Mix	2x	1 ml	1.6x
Primer premix	10x	200 ul	1.6x
ROX HIGH	50X	40 ul	1.6x

(1.24ml volume).

Table 43. qPCR Assay for Illumina Library Quantification.

Reagent	Starting Concentration	Volume Added	Final Concentration
Primer/dye master mix	1.6x	12.4 ul	1x
DNA standard or library dilution	na	4 ul	na
Nuclease-free water	na	3.6 ul	na

(20ul reactions).

APPENDIX C

GENOMIC SURVELLIANCE OF SARS-COV-2 VIRUS IN MONTANA 2020-2024

Introduction

The emergence of SARS-CoV-2 (SCoV2) virus in late 2019 initiated a global outbreak of covid19 respiratory disease, which was declared an official viral pandemic by the World Health Organization (WHO) from March 2020 to May 2023. During this time, there were 677 million cases reported, and 6.9 million deaths attributed to the disease, globally (Johns Hopkins Corona Virus Research Center). Much of this data was collected as part of a historic epidemiological surveillance effort by public health workers from around the world. Due to the recent advances in next generation sequencing technology, this surveillance effort also included tracking the emergence and transmission of new viral genome variants. This type of monitoring, called genomic surveillance, was critical to developing vaccine treatments and preventative public health policies aimed at slowing the spread of SCoV2 infection in our communities.

Genomic surveillance requires sequencing viral genomes sampled from infected individuals, and comparing the diversity of these genomes to what is observed globally through phylogenetic analysis. SCoV2 genome sequences determined by all academic, clinical, and government laboratories are deposited onto two publicly available repositories, called GISAID EpiCoV and NCBI GenBank. By comparing the sequence similarity of genomes in these databases, new variants become classified by their significance in the population and predicted risk of becoming a dominant circulating strain. These classifications include variants of high consequence (VHC), variants of concern (VOC), variants of interest (VOI), and variants under monitoring (VUM) (World Health Organization). Providing these classifications allows for easy communication of population genetics data to public health policy decision makers. However, there are many limitations to the data provided. Genomes within this database come from

different sampling, sequencing, and assembly methods. They are also paired with inconsistent demographic and geographic information. Additionally, the number of available sequences is biased towards large urban centers. These limitations make it difficult to compare samples across different regions. As such, many regional studies have since been conducted to showcase the emergence and evolution of SARS-CoV-2 in specific areas.

In areas where SARS-CoV-2 genomic surveillance has been assessed at a regional (country or state) level, there are clear differences in the patterns of SCoV2 variant emergence and transmission by geographic region [1-10]. The underlying causes of these differences are complex, but seem to include inconsistencies in travel to and from the region, population density within the region, testing and tracking capacity, access to preventative vaccination, and public health policies.

While genomic surveillance of SARS-CoV-2 is well described globally, data from rural geographic regions is particularly limited. Here, we provide the first comprehensive profile of SCoV2 genomic surveillance from the rural state of Montana for the entire collection period up to the date of writing this publication (December 2024). Our work includes presentation of genome variants by month, the whole genome nucleotide diversity by month, and clinical outcome indicators for a subset of this data from March to May of 2021. This work is meant to provide a historical snapshot of early emergence and tracking of SARS-CoV-2 in rural regions, describe how that has changed in the later years of the pandemic, and recommend future strategies for rapid implementation of genomic surveillance infrastructure in rural regions.

Materials and Methods

SARS-CoV-2 Genomic Data from GISAID.

A regional analysis of SARS-CoV2 (SCoV2) genomes collected from the state of Montana during January 1st, 2020 to December 31st, 2025 was completed using open-source data from the GISAID repository. Genomes obtained from GISAID were deposited by various academic research groups, hospitals, and public health facilities. Each group used a slightly different protocol for sample collection, processing, sequencing, and variant identification. These details are recorded in the GISAID submission. A ~4,000 sequence subset of the ~24,000 sequences in the dataset were deposited by our group and our collaborators. The following methods outline the protocols used to obtain viral genome sequences for these samples.

Clinical Specimen Collection.

SCoV2 variants were identified from nasal swab samples obtained by the Montana State University (MSU) CLIA lab and Bozeman Deaconess Health System Hospital. All samples tested positive for SCoV2 infection by RT-qPCR. Demographic metadata was variable by sample but included at least one of the following parameters - sex, age, onset condition, onset date, collection date, days from onset to collection, symptoms, cdc county code. All samples were deidentified and tracked on a secure server to protect patient privacy. Nasal swab samples are stored long term at -80 °C in the Clia lab facility at MSU.

High-throughput Viral RNA Extraction.

To evaluate hundreds of samples per week, we developed a high-throughput RNA extraction protocol using the Eppendorf epMotion liquid handler. A custom RNA extraction kit

was developed in collaboration with Promega to allowed for automated extraction of 96 samples at a time. Nasal swab samples were thawed at room temperature and 200 µl of suspension was transferred into a deep-well 96-well plate and combined with 220 µl of inactivation buffer. The inactivation reaction incubated at room temperature for 10 minutes. The plate was then placed onto the liquid handler and heat inactivated before beginning the custom RNA extraction procedure. After extraction, RNA samples were eluted into nuclease-free water for sequencing library preparation.

Illumina Sequencing using the ARTIC Protocol.

The NEB ARTIC protocol was used to prepare libraries for Illumina sequencing. This standard protocol begins with cDNA synthesis by reverse transcription reaction using the LunaScript RT kit. Amplification of cDNA was facilitated by the Q5 Hot Start High-Fidelity kit, followed by fragmentation and end-prep using the NEBNext Ultra II FS End Prep kit. Sample barcoding and adaptor ligation was completed with the NEBNext Ultra II Ligation kit. Final libraries were amplified with the NEBNext Library PCR kit. The concentrations of individually barcoded libraries were measured by PicoGreen fluorescence assays, and then pooled into a single sample with normalized concentration. The average fragment length and DNA concentration of the libraries were measured with Qubit and Bioanalyzer assays, respectively. Libraries were sequenced on the Illumina MiSeq machine at Montana State University.

Genome Assembly and Variant Identification.

We collaborated with the Good Lab at the University of Montana to develop a bioinformatic pipeline for genome assembly and variant identification, available at github.com/goodest-goodlab/UM-CoV-Seq. The main pipeline maps sequencing reads to the

SARS-CoV-2 reference genome, then uses GATK and iVar to call variants and create consensus sequences. These consensus sequences are then assigned to lineages using Pangolin and Nextclade.

Data Sharing.

All high-quality sequences were shared on NCBI GenBank and GISAID Open-Source Repositories, along with their associated variant identification, and subject demographic metadata.

Population Diversity Analysis.

The diversity index distribution was completed by performing a multiple sequence alignment in Mafft and calculating the diversity for each alignment in PopGenome. Alignments were grouped by the collection month of the sequenced sample.

General Regression Model for Clinical Associations.

Associations in the clinical data from a subset of samples collected in the Spring of 2021 were calculated by general regression modeling. All non-Ct variables were used to predict each of the Ct values. The “train” function from the R library “caret” was utilized to tune the model by picking the complexity parameters that were associated with optimal resampling statistics. The model “glmnet” was used for training, which fit generalized linear and similar models via penalized maximum likelihood. The statistics reported for each pairing were Estimate, Standard Error, t value, and p value.

Results and Discussion

In this study we examined data from the GISAID repository for the state of Montana from January 2020 to December 2024. This included a total of 24,530 samples, the majority of which were submitted during the year 2021. Samples were generally collected and processed using the same basic workflow (Fig 21A) where a patient, either showing symptoms of covid19 disease or with a known exposure to an infected individual, was tested for SCoV2 infection by a public health facility or clinical laboratory. Samples from patients testing positive for SCoV were processed for viral RNA extraction to isolate the genomic material. These genomes were then sequenced using next generation sequencing, and the resulting genome sequence compared to the ancestral Wuhan-Hu-1 strain (NC_045512.2) and other global circulating strains for phylogenetic classification and variant identification.

Global analysis of data held in GISAID can be difficult due to the wide variety of protocols and techniques used, especially during the early pandemic when protocols were still under development, for sample collection, processing, sequencing, and genome classification. GISAID submissions may specifically name protocol features such ‘specimen source’, ‘sequencing technology’, and ‘assembly method’ but are not required to do so. This is why smaller studies of regions that may have been working together to develop similar surveillance techniques provide the highest quality data for interpretation.

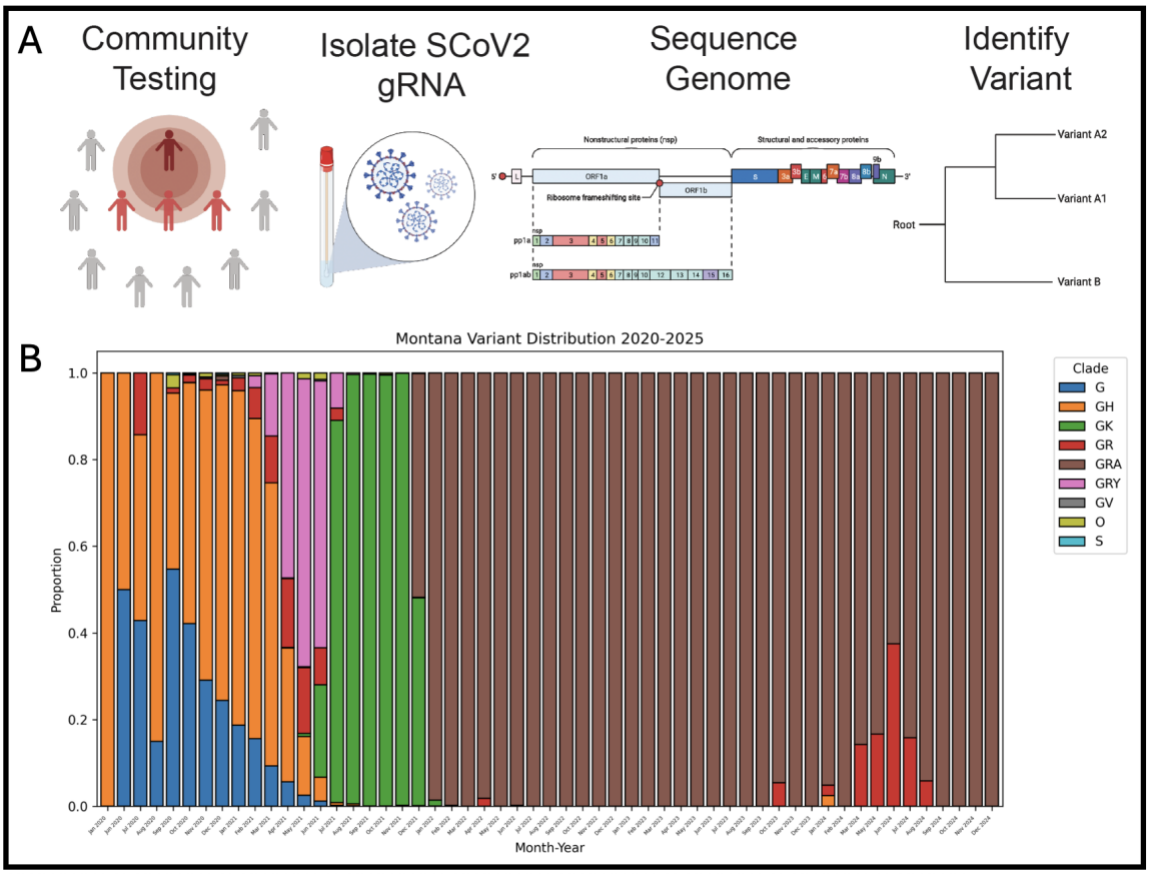


Figure 21. Genomic Surveillance of SARS-CoV-2 (SCoV2) Virus in Montana from 2020-2025. (A) The sample collection and processing workflow included community testing by public health authorities and clinical laboratories, followed by isolation of SCoV2 genome RNA from positive samples, then genomic sequencing and variant identification in comparison to the global dataset and ancestral strain. (B) SCoV2 variant distribution from all samples (n = 24,530) published to GISAID from January 2020 to December of 2024, showing the proportion of each unique variant clade in the population by month.

In our study, looking at samples from only the state of Montana, there were large inconsistencies in the next generation sequencing platform used to obtain genomic data. The majority of the samples were sequenced using an Oxford Nanopore platform (n = 10,354), and the rest were sequenced with either an Illumina platform (n = 8,613) or did not specify (n = 5,563). Additionally, GISAID does not specify the method or quantification used to test an individual for positive infection, which generally was by RT-qPCR testing at the start of the

pandemic and shifted to antigen testing near the end. Another large inconsistency in the data is the number of samples collected per year. The majority of our data is skewed to the early years of the pandemic, including 2020 (n = 7,487), 2021 (n = 10,499), and 2022 (n = 6,011). There was a sharp drop off in 2023 (n=341) and 2024 (n=191), most likely due to increased vaccination, reduced case numbers, and the widespread use of at-home testing kits that does not require any sort of reporting of positive cases.

The distribution of SCoV2 variants within our dataset (Fig 21B) shows a heterogeneous viral population at the start of the tracking period, with many different clades circulating at detectable levels in the population at once, that becomes more homogenous over time. There are a few clear waves of emergence illustrated in the distribution, where a new variant becomes dominant in the population for a period of time. Most interestingly, perhaps, is a re-emergence event in the summer of 2024 by the GR clade, after being undetectable in the population for over two years, during a period of long-standing dominance by GRA clade. Diversity within the variant distribution quantified by a whole genome nucleotide diversity index calculated for each collection month in the dataset (Fig 22A).

The diversity index distribution shows a general trend in increased population diversity over time, although the variant distribution shows fewer unique clades present. Additionally, there was large spike in nucleotide diversity in the population in the winter of 2023, just before the re-emergence of the GR clade in the spring of 2024. Together these trends in these distributions show that the SCoV2 population in Montana began with many variants whose genomes are more closely related to each other than to the current dominant variant (GRA)

which contains highly diverse genomes within a single clade. However, this trend again could be skewed by significant differences in the number of samples available for analysis in each year.

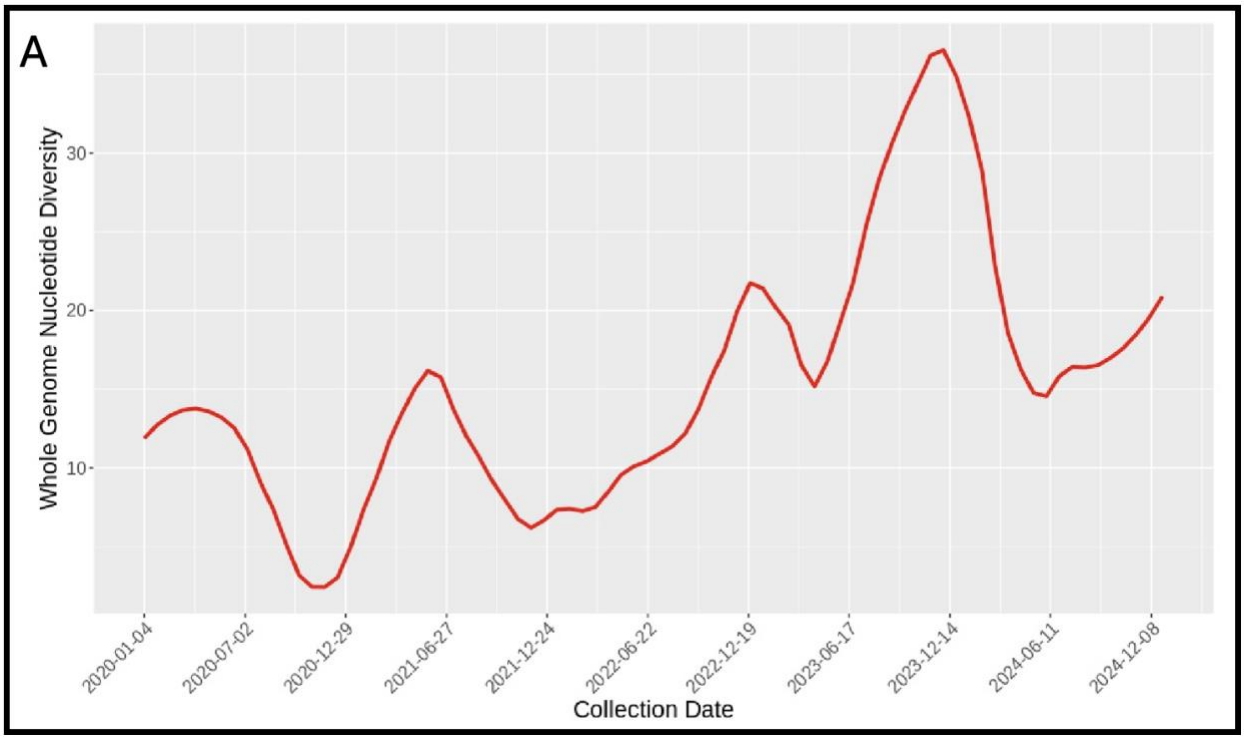


Figure 22. SARS-CoV-2 Whole Genome Nucleotide Diversity from 2020-2024 in Montana. (A) SCoV2 diversity index distribution from all samples (n = 24,530) published to GISAID from January 2020 to December of 2024, showing the whole nucleotide genome diversity within the population by month.

Most community testing and genomic surveillance took place during 2021, when many people were still unvaccinated against SCoV2 and at home testing had not yet dominated the diagnosis market. This was also a period where many minor variants of concern were still in circulation before the first major wave of variant dominance with the Delta strain in May 2021. During this time, we collected a small subset of samples from 91 individuals in Bozeman, MT who provided valuable demographic and disease presentation information (Fig 23). For each sample we collected viral load quantity measured by the cycle threshold (*Ct*) value of an RT-qPCR test. These quantities were measured for three important viral genes, nucleocapsid

phosphoprotein (N) gene, open-reading frame 1ab (Orf1ab), and the spike protein (S) gene. Each sample was also correlated to a list of demographic and clinical information including cdc county density code, patient sex, patient age, onset condition (if a person was symptomatic or exposed to an infected individual, onset date, collection date, days from onset to collection, symptoms, and viral variant ID. The raw relationship between gene specific *Ct* value and patient sex (Fig 23A), patient age (Fig 23B), and onset condition (Fi 23C) were plotted below.

To determine the significance of these associations, as well as the other demographic data not shown, three general regression models were developed (Fig 23D) with the N gene, Orf1ab, and S genes as individual independent variables. The general regression models showed a significant negative association between N gene quantity and prevalence of the B.1.1.7 strain in the population ($p = 0.0142$), Orf1ab quantity and cdc population density county code ($p= 0.0388$), as well as S gene quantity and patients within the 11 to 44 age range ($p=0.0689$). Additionally, the models showed significant positive association between S gene quantity and the number of days between symptom onset and sample collection ($p = 0.0054$), as well as patients self-reporting sex as Female ($p = 4.01x10^{-5}$). While this data is representative of a small subset of covid19 patients (91) it provides a critical gap in the literature for disease presentation in rural areas, where genomic surveillance and general epidemiological data for the covid19 pandemic is very limited. The associations determined here should be compared to similar clinical and demographic data from other regional studies, to understand just how widely patterns of SCoV2 infection varied by population density.

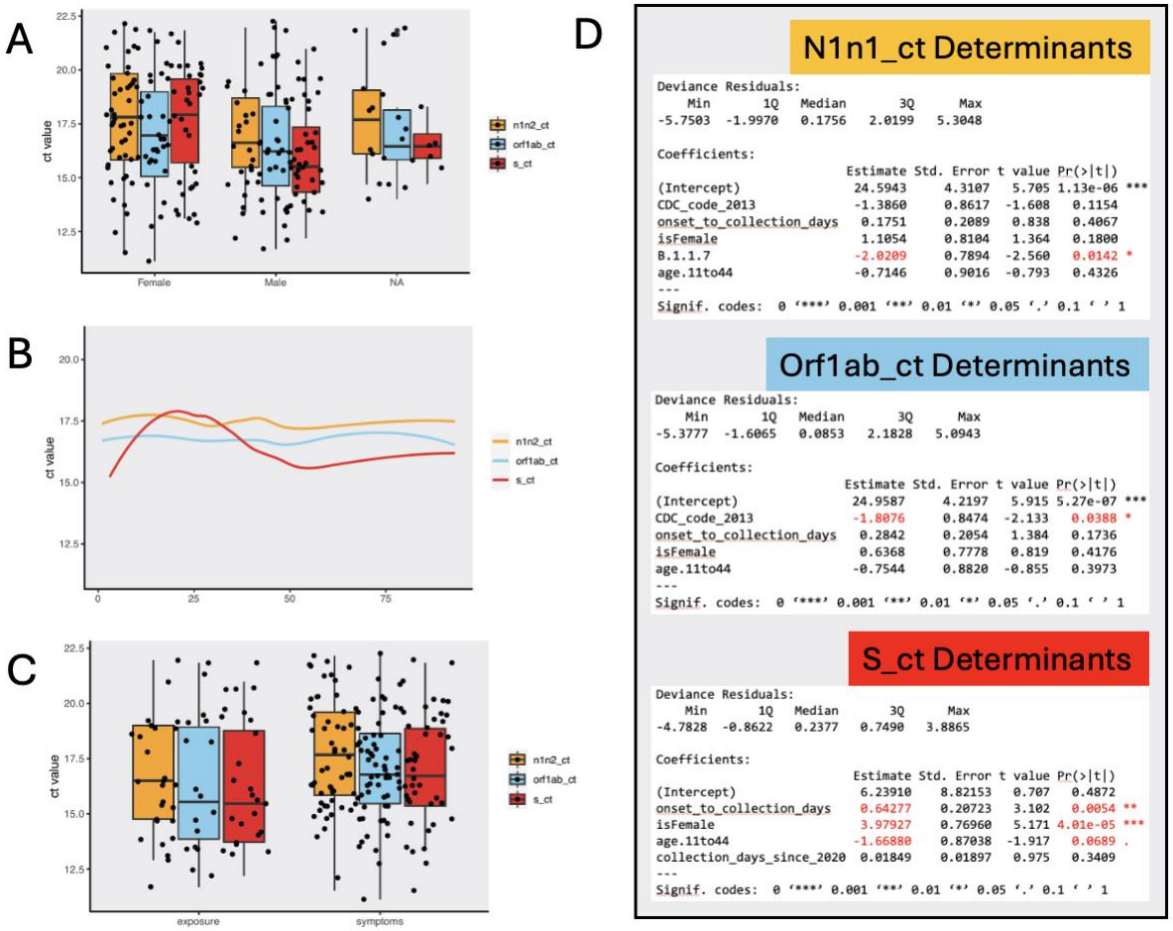


Figure 23. Clinical Presentation of SARS-CoV-2 Infection from Bozeman, MT in Spring 2021. Subset of samples (n = 91) collected from March to May of 2021 included demographic and disease presentation data, along with the standard RT-qPCR cycle threshold (*Ct*) value for three genes of interest: nucleocapsid phosphoprotein (N) gene (n1n1_ct, orange), open-reading frame 1ab (Orf1ab) (orf1ab_ct, blue), and spike protein (S) gene (s_ct, red). The relationship between *Ct* value and self-reported demographics like (A) biological sex (B) age, and (C) testing rationale shown as raw data plotted without statistical significance. (D) General regression models were generated for three datasets to determine the impact of selected demographic variables (cdc county code, sex, age, onset condition, onset date, collection date, days from onset to collection, symptoms, viral variant ID) on gene specific *Ct* values (N gene, Orf1ab, S gene).

Conclusion

The covid19 pandemic mounted an unprecedented effort in the mobilization of public health prevention and policy measures. One such measure was widespread genomic surveillance

of the emerging SCoV2 virus. The advent of next generation sequencing technology made such an effort possible, but it's application was not equally distributed throughout geographic regions. Rural regions faced large funding and infrastructure barriers to collecting genomic surveillance data during the covid19 pandemic, due to a lack of medical schools, hospitals, and research facilities with the required staff and equipment to undertake such data collection. Here we presented a snapshot of standard genomic surveillance data collected from the rural state of Montana during this time. We capture a few interesting patterns in variant emergence and diversity patterns, as well as associations between such variants and clinical disease presentation of infected individuals from this area. On it's own, this data is meant to serve as a historical record of how a rural area was able to mount an ambitious genomic surveillance program with limited resources. In combination with similar data collected from other rural regions, and in comparison, to data from urban centers, we may begin to use this information to understand how SCoV2 emergence and evolution patterns may have varied by population density during the covid19 pandemic. All together this work may help to aid in the preparedness for future pandemic events that are surely to affect our global community once more.

References Cited

1. Deng, X. *et al.* Genomic surveillance reveals multiple introductions of SARS-CoV-2 into Northern California. *Science* **369**, 582–587 (2020).

2. Butera, Y. *et al.* Genomic sequencing of SARS-CoV-2 in Rwanda reveals the importance of incoming travelers on lineage diversity. *Nat. Commun.* **12**, 5705 (2021).

3. Wilkinson, E. *et al.* A year of genomic surveillance reveals how the SARS-CoV-2 pandemic unfolded in Africa. *Science* **374**, 423–431 (2021).

4. Ozer, E. A. *et al.* Multiple expansions of globally uncommon SARS-CoV-2 lineages in Nigeria. *Nat. Commun.* **13**, 688 (2022).

5. Santiago, G. A. *et al.* Genomic surveillance of SARS-CoV-2 in Puerto Rico enabled early detection and tracking of variants. *Commun. Med.* **2**, 100 (2022).

6. Giovanetti, M. *et al.* Genomic epidemiology of the SARS-CoV-2 epidemic in Brazil. *Nat Microbiol* (2022) doi:10.1038/s41564-022-01191-z.

7. Arantes, I. *et al.* Comparative epidemic expansion of SARS-CoV-2 variants Delta and Omicron in the Brazilian State of Amazonas. *Nat. Commun.* **14**, 2048 (2023).

8. Koné, A. *et al.* Dynamics of SARS-CoV-2 variants characterized during different COVID-19 waves in Mali. *IJID Reg* **6**, 24–28 (2023).

9. Oltean, H. N. *et al.* Sentinel Surveillance System Implementation and Evaluation for SARS-CoV-2 Genomic Data, Washington, USA, 2020-2021. *Emerg. Infect. Dis.* **29**, 242–251 (2023).

10. Fokam, J. *et al.* Genomic surveillance of SARS-CoV-2 reveals highest severity and mortality of delta over other variants: evidence from Cameroon. *Sci. Rep.* **13**, 21654 (2023).

CUMULATIVE REFERENCES CITED

Abate, A. R., Chen, C.-H., Agresti, J. J. & Weitz, D. A. Beating Poisson encapsulation statistics using close-packed ordering. *Lab Chip* **9**, 2628–2631 (2009).

Acheson, N. Fundamentals of molecular virology 2nd edition - NLM Catalog - NCBI. https://www.ncbi.nlm.nih.gov/nlmcatalog/101551557 (2011).

Acklin, J. A. *et al.* Immunological landscape of human lymphoid explants during measles virus infection. *JCI Insight* **9**, e172261 (2024).

Adams, S. K., Ducharme, G. E. & Loveday, E. K. All the single cells: if you like it then you should put some virus on it. *J. Virol.* **98**, e0127323 (2024).

Agresti, J. J. *et al.* Ultrahigh-throughput screening in drop-based microfluidics for directed evolution. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 4004–4009 (2010).

Akpinar, F., Inankur, B. & Yin, J. Spatial-Temporal Patterns of Viral Amplification and Interference Initiated by a Single Infected Cell. *J. Virol.* **90**, 7552–7566 (2016).

Akpinar, F., Timm, A. & Yin, J. High-Throughput Single-Cell Kinetics of Virus Infections in the Presence of Defective Interfering Particles. *J. Virol.* **90**, 1599–1612 (2016).

Alnaji, F. G. *et al.* Sequencing Framework for the Sensitive Detection and Precise Mapping of Defective Interfering Particle-Associated Deletions across Influenza A and B Viruses. *J. Virol.* **93**, (2019).

Altschuler, S. J. & Wu, L. F. Cellular heterogeneity: do differences make a difference? *Cell* **141**, 559–563 (2010).

Alvarez, M. J., Vila-Ortiz, G. J., Salibe, M. C., Podhajcer, O. L. & Pitossi, F. J. Model based analysis of real-time PCR data from DNA binding dye protocols. *BMC Bioinformatics* **8**, 85 (2007).

Anderson, B. D., Nakamura, T., Russell, S. J. & Peng, K.-W. High CD46 receptor density determines preferential killing of tumor cells by oncolytic measles virus. *Cancer Res.* **64**, 4919–4926 (2004).

Andino, R. & Domingo, E. Viral quasispecies. *Virology* **479–480**, 46–51 (2015).

Arantes, I. *et al.* Comparative epidemic expansion of SARS-CoV-2 variants Delta and Omicron in the Brazilian State of Amazonas. *Nat. Commun.* **14**, 2048 (2023).

Aref, S., Bailey, K. & Fielding, A. Measles to the Rescue: A Review of Oncolytic Measles Virus. *Viruses* **8**, 294 (2016).

Bacsik, D. J. *et al.* Influenza virus transcription and progeny production are poorly correlated in

single cells. *bioRxiv* 2022.08.30.505828 (2022) doi:10.1101/2022.08.30.505828.

Baltimore, D. Viral genetic systems. *Trans. N. Y. Acad. Sci.* **33**, 327–332 (1971).

Beer, N. R. *et al.* On-chip single-copy real-time reverse-transcription PCR in isolated picoliter droplets. *Anal. Chem.* **80**, 1854–1858 (2008).

Beer, N. R. *et al.* On-chip, real-time, single-copy polymerase chain reaction in picoliter droplets. *Anal. Chem.* **79**, 8471–8475 (2007).

Bellini, W. J. & Rota, P. A. Genetic diversity of wild-type measles viruses: implications for global measles elimination programs. *Emerg. Infect. Dis.* **4**, 29–35 (1998).

Bhat, T., Cao, A. & Yin, J. Virus-like Particles: Measures and Biological Functions. *Viruses* **14**, (2022).

Bost, P. & Drayman, N. Dissecting viral infections, one cell at a time, by single-cell technologies. *Microbes Infect.* **26**, 105268 (2024).

Brooke, C. B. *et al.* Most influenza a virions fail to express at least one essential viral protein. *J. Virol.* **87**, 3155–3162 (2013).

Brooke, C. B. Population Diversity and Collective Interactions during Influenza Virus Infection. *J. Virol.* **91**, (2017).

Brooke, C. B. Population Diversity and Collective Interactions during Influenza Virus Infection. *J. Virol.* **91**, (2017).

Budnik, B., Levy, E., Harmange, G. & Slavov, N. SCoPE-MS: mass spectrometry of single mammalian cells quantifies proteome heterogeneity during cell differentiation. *Genome Biol.* **19**, 161 (2018).

Burnet, F. M. A Method for the Study of Bacteriophage Multiplication in Broth. *Br. J. Exp. Pathol.* **10**, 109 (1929).
Butera, Y. *et al.* Genomic sequencing of SARS-CoV-2 in Rwanda reveals the importance of incoming travelers on lineage diversity. *Nat. Commun.* **12**, 5705 (2021).

Butler, D. *et al.* Shotgun transcriptome, spatial omics, and isothermal profiling of SARS-CoV-2 infection reveals unique host responses, viral diversification, and drug interactions. *Nat. Commun.* **12**, 1660 (2021).

Carlo, D. D., Irimia, D., Tompkins, R. G. & Toner, M. Continuous inertial focusing, ordering, and separation of particles in microchannels. *Proceedings of the National Academy of Sciences* **104**, 18892–18897 (2007).

Chang, C., Bianco, S., Acevedo, A., Tang, C. & Andino, R. Genetic interactions shaping evolutionary trajectories in an RNA virus population. *bioRxiv* 2020.01.16.908129 (2020) doi:10.1101/2020.01.16.908129.

Chattopadhyay, P. K., Roederer, M. & Bolton, D. L. A deadly dance: the choreography of host-pathogen interactions, as revealed by single-cell technologies. *Nat. Commun.* **9**, 4638 (2018).

Chen, K.-Y. *et al.* High-throughput droplet-based analysis of influenza A virus genetic reassortment by single-virus RNA sequencing. *Proc. Natl. Acad. Sci. U. S. A.* **120**, e2211098120 (2023).

Chu, C., Lugovtsev, V., Golding, H., Betenbaugh, M. & Shiloach, J. Conversion of MDCK cell line to suspension culture by transfecting with human siat7e gene and its application for influenza virus production. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 14802–14807 (2009).

Chu, C., Lugovtsev, V., Lewis, A., Betenbaugh, M. & Shiloach, J. Production and antigenic properties of influenza virus from suspension MDCK-siat7e cells in a bench-scale bioreactor. *Vaccine* **28**, 7193–7201 (2010).

Ciuffi, A., Rato, S. & Telenti, A. Single-Cell Genomics for Virology. *Viruses* **8**, (2016).

Clark, I. C. *et al.* Microfluidics-free single-cell genomics with templated emulsification. *Nat. Biotechnol.* **41**, 1557–1566 (2023).

Clementi, M. *et al.* Quantitative PCR and RT-PCR in virology. *PCR methods and applications* **2**, 191–191 (1993).

Combe, M., Garijo, R., Geller, R., Cuevas, J. M. & Sanjuán, R. Single-Cell Analysis of RNA Virus Infection Identifies Multiple Genetically Diverse Viral Genomes within Single Infectious Units. *Cell Host Microbe* **18**, 424–432 (2015).

Condack, C., Grivel, J.-C., Devaux, P., Margolis, L. & Cattaneo, R. Measles virus vaccine attenuation: suboptimal infection of lymphatic tissue and tropism alteration. *J. Infect. Dis.* **196**, 541–549 (2007).

Corder, B. N., Bullard, B. L., Poland, G. A. & Weaver, E. A. A Decade in Review: A Systematic Review of Universal Influenza Vaccines in Clinical Trials during the 2010 Decade. *Viruses* **12**, (2020).

Crick, F. Central dogma of molecular biology. *Nature* **227**, 561–563 (1970).

Cristinelli, S. & Ciuffi, A. The use of single-cell RNA-Seq to understand virus-host interactions. *Curr. Opin. Virol.* **29**, 39–50 (2018).

Cummings, K. W., Levy, D. N. & Wodarz, D. Increased burst size in multiply infected cells can alter basic virus dynamics. *Biol. Direct* **7**, 16 (2012).

Davis, A. R., Hiti, A. L. & Nayak, D. P. Influenza defective interfering viral RNA is formed by internal deletion of genomic RNA. *Proc. Natl. Acad. Sci. U. S. A.* **77**, 215–219 (1980).

Delbrück, M. The Burst Size Distribution in the Growth of Bacterial Viruses (Bacteriophages). *J. Bacteriol.* **50**, 131–135 (1945).

Delley, C. L. & Abate, A. R. Modular barcode beads for microfluidic single cell genomics. *Sci. Rep.* **11**, 10857 (2021).

DeLong, J. P. *et al.* Towards an integrative view of virus phenotypes. *Nat. Rev. Microbiol.* (2021) doi:10.1038/s41579-021-00612-w.
Deng, X. *et al.* Genomic surveillance reveals multiple introductions of SARS-CoV-2 into Northern California. *Science* **369**, 582–587 (2020).

Devaux, P., von Messling, V., Songsungthong, W., Springfeld, C. & Cattaneo, R. Tyrosine 110 in the measles virus phosphoprotein is required to block STAT1 phosphorylation. *Virology* **360**, 72–83 (2007).

Devaux, P., Priniski, L. & Cattaneo, R. The measles virus phosphoprotein interacts with the linker domain of STAT1. *Virology* **444**, 250–256 (2013).

Dixit, A. *et al.* Perturb-seq: Dissecting molecular circuits with scalable single-cell RNA profiling of pooled genetic screens. *Cell* **167**, 1853-1866.e17 (2016).

Dolan, P. T., Whitfield, Z. J. & Andino, R. Mapping the Evolutionary Potential of RNA Viruses. *Cell Host Microbe* **23**, 435–446 (2018).

Dolan, P. T., Whitfield, Z. J. & Andino, R. Mechanisms and concepts in RNA virus population dynamics and evolution. *Annu. Rev. Virol.* **5**, 69–92 (2018).

Domingo, E., Sheldon, J. & Perales, C. Viral quasispecies evolution. *Microbiol. Mol. Biol. Rev.* **76**, 159–216 (2012).

Domingo, E. *et al.* Basic concepts in RNA virus evolution. *FASEB J.* **10**, 859–864 (1996).

Domingo, E. *RNA Genetics: Volume III: Variability of RNA Genomes*. (CRC Press, London, England, 2020). doi:10.1201/9781351076449.

Domingo, E. Viruses at the edge of adaptation. *Virology* **270**, 251–253 (2000).

Domingo, E. & Holland, J. J. RNA virus mutations and fitness for survival. *Annu. Rev. Microbiol.* **51**, 151–178 (1997).

Drake, J. W. Rates of spontaneous mutation among RNA viruses. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 4171–4175 (1993).

Drayman, N., Patel, P., Vistain, L. & Tay, S. HSV-1 single-cell analysis reveals the activation of anti-viral and developmental programs in distinct sub-populations. *Elife* **8**, (2019).

Duffy, D. C., McDonald, J. C., Schueller, O. J. & Whitesides, G. M. Rapid prototyping of microfluidic systems in poly(dimethylsiloxane). *Anal. Chem.* **70**, 4974–4984 (1998).

Duffy, S., Shackelton, L. A. & Holmes, E. C. Rates of evolutionary change in viruses: patterns and determinants. *Nat. Rev. Genet.* **9**, 267–276 (2008).

Duhaut, S. D. & McCauley, J. W. Defective RNAs inhibit the assembly of influenza virus genome segments in a segment-specific manner. *Virology* **216**, 326–337 (1996).

Durnell, L. A. *et al.* Interferon-independent processes constrain measles virus cell-to-cell spread in primary human airway epithelial cells. *Microbiol. Spectr.* e0136123 (2023).

Dutta, R. N., Rouzine, I. M., Smith, S. D., Wilke, C. O. & Novella, I. S. Rapid adaptive amplification of preexisting variation in an RNA virus. *J. Virol.* **82**, 4354–4362 (2008).

Eastburn, D. J., Sciambi, A. & Abate, A. R. Ultrahigh-throughput Mammalian single-cell reverse-transcriptase polymerase chain reaction in microfluidic drops. *Anal. Chem.* **85**, 8016–8021 (2013).

Eigen, M. Viral quasispecies. *Scientific American* **269**, 42–49 (1993).

Engeland, C. E. & Ungerechts, G. Measles virus as an oncolytic immunotherapy. *Cancers (Basel)* **13**, 544 (2021).

Finkelman, B. S. *et al.* Global patterns in seasonal activity of influenza A/H3N2, A/H1N1, and B from 1997 to 2005: viral coexistence and latitudinal gradients. *PLoS One* **2**, e1296 (2007).

Fischer, A. E. *et al.* A high-throughput drop microfluidic system for virus culture and analysis. *J. Virol. Methods* **213**, 111–117 (2015).

Fokam, J. *et al.* Genomic surveillance of SARS-CoV-2 reveals highest severity and mortality of delta over other variants: evidence from Cameroon. *Sci. Rep.* **13**, 21654 (2023).

Frishberg, A. *et al.* Cell composition analysis of bulk genomics using single-cell data. *Nat. Methods* **16**, 327–332 (2019).

Galanis, E. *et al.* Oncolytic measles virus expressing the sodium iodide symporter to treat drug-resistant ovarian cancer. *Cancer Res.* **75**, 22–30 (2015).

Gallagher, M. E., Brooke, C. B., Ke, R. & Koelle, K. Causes and Consequences of Spatial Within-Host Viral Spread. *Viruses* **10**, (2018).

Gérard, A. *et al.* High-throughput single-cell activity-based screening and sequencing of antibodies using droplet microfluidics. *Nat. Biotechnol.* (2020) doi:10.1038/s41587-020-0466-7.

Gevertz, J. L., Dunn, S. M. & Roth, C. M. Mathematical model of real-time PCR kinetics. *Biotechnol. Bioeng.* **92**, 346–355 (2005).

Ghedin, E. *et al.* Large-scale sequencing of human influenza reveals the dynamic nature of viral genome evolution. *Nature* **437**, 1162–1166 (2005).

Gini, C. Measurement of Inequality of Incomes. *Econ J* **31**, 124–125 (1921).

Gini, C. *Variabilità e Mutabilità*. (1912).

Giovanetti, M. *et al.* Genomic epidemiology of the SARS-CoV-2 epidemic in Brazil. *Nat Microbiol* (2022) doi:10.1038/s41564-022-01191-z.

Griffin, D. E., Lin, W.-H. & Pan, C.-H. Measles virus, immune control, and persistence. *FEMS Microbiol. Rev.* **36**, 649–662 (2012).

Gross, A. *et al.* Technologies for Single-Cell Isolation. *Int. J. Mol. Sci.* **16**, 16897–16919 (2015).

Guo, F. *et al.* Single-Cell Virology: On-Chip Investigation of Viral Infection Dynamics. *Cell Rep.* **21**, 1692–1704 (2017).

Guo, M. T., Rotem, A., Heyman, J. A. & Weitz, D. A. Droplet microfluidics for high-throughput biological assays. *Lab Chip* **12**, 2146–2155 (2012).

Guo, M. T., Rotem, A., Heyman, J. A. & Weitz, D. A. Droplet microfluidics for high-throughput biological assays. *Lab Chip* **12**, 2146–2155 (2012).

Hajji, I. *et al.* Droplet microfluidic platform for fast and continuous-flow RT-qPCR analysis devoted to cancer diagnosis application. *Sens. Actuators B Chem.* **303**, 127171 (2020).

Hataye, J. M. *et al.* Principles Governing Establishment versus Collapse of HIV-1 Cellular Spread. *Cell Host Microbe* **26**, 748-763.e20 (2019).

Hatch, A. C., Ray, T., Lintecum, K. & Youngbull, C. Continuous flow real-time PCR device

using multi-channel fluorescence excitation and detection. *Lab Chip* **14**, 562–568 (2014).

Hatch, A. C. *et al.* 1-Million droplet array with wide-field fluorescence imaging for digital PCR. *Lab Chip* **11**, 3838 (2011).

Hatch, A. C. *et al.* 1-Million droplet array with wide-field fluorescence imaging for digital PCR. *Lab Chip* **11**, 3838–3845 (2011).

Hayden, E. J., Ferrada, E. & Wagner, A. Cryptic genetic variation promotes rapid evolutionary adaptation in an RNA enzyme. *Nature* **474**, 92–95 (2011).

Hayes, C. J. & Dalton, T. M. Microfluidic droplet-based PCR instrumentation for high-throughput gene expression profiling and biomarker discovery. *Biomol Detect Quantif* **4**, 22–32 (2015).

Heldt, F. S., Kupke, S. Y., Dorl, S., Reichl, U. & Frensing, T. Single-cell analysis and stochastic modelling unveil large cell-to-cell variability in influenza A virus infection. *Nat. Commun.* **6**, 8938 (2015).

Heldt, F. S., Frensing, T. & Reichl, U. Modeling the intracellular dynamics of influenza virus replication to understand the control of viral RNA synthesis. *J. Virol.* **86**, 7806–7817 (2012).

Hindson, B. J. *et al.* High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. *Anal. Chem.* **83**, 8604–8610 (2011).

Hippee, C. E. *et al.* Measles virus exits human airway epithelia within dislodged metabolically active infectious centers. *PLoS Pathog.* **17**, e1009458 (2021).

Hippee, C. E. *et al.* Epithelial-to-mesenchymal transition and live cell extrusion contribute to measles virus release from human airway epithelia. *J. Virol.* **99**, e0122024 (2025).

Hockman, M. R., Phipps, K. L., Holmes, K. E. & Lowen, A. C. A method for the unbiased quantification of reassortment in segmented viruses. *J. Virol. Methods* **280**, 113878 (2020).

Hoffmann, E., Stech, J., Guan, Y., Webster, R. G. & Perez, D. R. Universal primer set for the full-length amplification of all influenza A viruses. *Arch. Virol.* **146**, 2275–2289 (2001).

Holland, J. *et al.* Rapid evolution of RNA genomes. *Science* **215**, 1577–1585 (1982).

Holland, J. J., De La Torre, J. C. & Steinhauer, D. A. RNA virus populations as quasispecies. *Curr. Top. Microbiol. Immunol.* **176**, 1–20 (1992).

Holmes, M., Zhang, F. & Bieniasz, P. D. Single-Cell and Single-Cycle Analysis of HIV-1 Replication. *PLoS Pathog.* **11**, e1004961 (2015).

Holtze, C. *et al.* Biocompatible surfactants for water-in-fluorocarbon emulsions. *Lab Chip* **8**, 1632–1639 (2008).

Huang, A. S. & Baltimore, D. Defective viral particles and viral disease processes. *Nature* **226**, 325–327 (1970).

Jacobs, N. T. *et al.* Incomplete influenza A virus genomes occur frequently but are readily complemented during localized viral spread. *Nat. Commun.* **10**, 3526 (2019).

Jiang, S. *et al.* Combined protein and nucleic acid imaging reveals virus-dependent B cell and macrophage immunosuppression of tissue microenvironments. *Immunity* **55**, 1118-1134.e8 (2022).

Jones, J. E., Le Sage, V. & Lakdawala, S. S. Viral and host heterogeneity and their effects on the viral life cycle. *Nat. Rev. Microbiol.* **19**, 272–282 (2021).

Kaminski, T. S., Scheler, O. & Garstecki, P. Droplet microfluidics for microbiology: techniques, applications and challenges. *Lab Chip* **16**, 2168–2187 (2016).

Kasibhatla, S. M., Vaidya, S. R., Kale, M. M. & Kulkarni-Kale, U. Phylogenomics and evolution of measles virus. in *Phylogenomics* 391–413 (Elsevier, 2024).

Ke, R., Lewin, S. R., Elliott, J. H. & Perelson, A. S. Modeling the Effects of Vorinostat In Vivo Reveals both Transient and Delayed HIV Transcriptional Activation and Minimal Killing of Latently Infected Cells. *PLoS Pathog.* **11**, e1005237 (2015).

Kemler, I., Ennis, M. K., Neuhauser, C. M. & Dingli, D. In vivo imaging of oncolytic measles virus propagation with single-cell resolution. *Mol. Ther. Oncolytics* **12**, 68–78 (2019).

Kiss, M. M. *et al.* High-throughput quantitative polymerase chain reaction in picoliter droplets. *Anal. Chem.* **80**, 8975–8981 (2008).

Klein, A. M. *et al.* Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* **161**, 1187–1201 (2015).

Koné, A. *et al.* Dynamics of SARS-CoV-2 variants characterized during different COVID-19 waves in Mali. *IJID Reg* **6**, 24–28 (2023).

Köster, S. *et al.* Drop-based microfluidic devices for encapsulation of single cells. *Lab Chip* **8**, 1110–1115 (2008).

Krammer, F. *et al.* Influenza. *Nat Rev Dis Primers* **4**, 3 (2018).

Krokhine, S., Torabi, H., Doostmohammadi, A. & Rezai, P. Conventional and microfluidic methods for airborne virus isolation and detection. *Colloids Surf. B Biointerfaces* **206**, 111962 (2021).

Kühne, M., Brown, D. W. G. & Jin, L. Genetic variability of measles virus in acute and persistent infections. *Infect. Genet. Evol.* **6**, 269–276 (2006).

Kupke, S. Y. *et al.* Single-Cell Analysis Uncovers a Vast Diversity in Intracellular Viral Defective Interfering RNA Content Affecting the Large Cell-to-Cell Heterogeneity in Influenza A Virus Replication. *Viruses* **12**, (2020).

Kupke, S. Y. *et al.* Single-Cell Analysis Uncovers a Vast Diversity in Intracellular Viral Defective Interfering RNA Content Affecting the Large Cell-to-Cell Heterogeneity in Influenza A Virus Replication. *Viruses* **12**, (2020).

Kurokawa, C. *et al.* Constitutive Interferon Pathway Activation in Tumors as an Efficacy Determinant Following Oncolytic Virotherapy. *J. Natl. Cancer Inst.* **110**, 1123–1132 (2018).

Lauring, A. S. & Andino, R. Quasispecies theory and the behavior of RNA viruses. *PLoS Pathog.* **6**, e1001005 (2010).

Lee, J., Hyeon, D. Y. & Hwang, D. Single-cell multiomics: technologies and data analysis methods. *Exp. Mol. Med.* **52**, 1428–1442 (2020).

Lehnert, T. & Gijs, M. A. M. Microfluidic systems for infectious disease diagnostics. *Lab Chip* **24**, 1441–1493 (2024).

Liao, M. *et al.* Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. *Nat. Med.* **26**, 842–844 (2020).

Lievens, A., Van Aelst, S., Van den Bulcke, M. & Goetghebeur, E. Enhanced analysis of real-time PCR data by using a variable efficiency model: FPK-PCR. *Nucleic Acids Res.* **40**, e10 (2012).

Lin, J. *et al.* Ultra-sensitive digital quantification of proteins and mRNA in single cells. *Nat. Commun.* **10**, 3544 (2019).

Liu, W., He, H. & Zheng, S.-Y. Microfluidics in Single-Cell Virology: Technologies and Applications. *Trends Biotechnol.* (2020) doi:10.1016/j.tibtech.2020.04.010.

Liu, W. & Saint, D. A. Validation of a quantitative method for real time PCR kinetics. *Biochem. Biophys. Res. Commun.* **294**, 347–353 (2002).

Liu, W., He, H. & Zheng, S.-Y. Microfluidics in single-cell virology: Technologies and

applications. *Trends Biotechnol.* **38**, 1360–1372 (2020).

Loveday, E. K., Sanchez, H. S., Thomas, M. M. & Chang, C. B. Single-Cell Infection of Influenza A Virus Using Drop-Based Microfluidics. *Microbiol Spectr* e0099322 (2022).

Loveday, E. K., Zath, G. K., Bikos, D. A., Jay, Z. J. & Chang, C. B. Screening of Additive Formulations Enables Off-Chip Drop Reverse Transcription Quantitative Polymerase Chain Reaction of Single Influenza A Virus Genomes. *Anal. Chem.* **93**, 4365–4373 (2021).

Mackay, I. M., Arden, K. E. & Nitsche, A. Real-time PCR in virology. *Nucleic Acids Res.* **30**, 1292–1305 (2002).

Macosko, E. Z. *et al.* Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015).

Martin, B. E., Harris, J. D., Sun, J. & Koelle, K. Cellular co-infection can modulate the efficiency of influenza A virus production and shape the interferon response. *PLoS* (2020).

Mattiazzi Usaj, M., Yeung, C. H. L., Friesen, H., Boone, C. & Andrews, B. J. Single-cell image analysis to explore cell-to-cell heterogeneity in isogenic populations. *Cell Syst* **12**, 608–621 (2021).

Matuła, K., Rivello, F. & Huck, W. T. S. Single-Cell Analysis Using Droplet Microfluidics. *Adv Biosyst* **4**, e1900188 (2020).

Mazutis, L. *et al.* Single-cell analysis and sorting using droplet-based microfluidics. *Nat. Protoc.* **8**, 870–891 (2013).

Mazutis, L. *et al.* Single-cell analysis and sorting using droplet-based microfluidics. *Nature protocols* **8**, 870–891 (2013).

McCrone, J. T. *et al.* Stochastic processes constrain the within and between host evolution of influenza virus. *Elife* **7**, (2018).

Mimitou, E. P. *et al.* Multiplexed detection of proteins, transcriptomes, clonotypes and CRISPR perturbations in single cells. *Nat. Methods* **16**, 409–412 (2019).

Moffitt, J. R., Lundberg, E. & Heyn, H. The emerging landscape of spatial profiling technologies. *Nat. Rev. Genet.* **23**, 741–759 (2022).

Molinari, N.-A. M. *et al.* The annual impact of seasonal influenza in the US: measuring disease burden and costs. *Vaccine* **25**, 5086–5096 (2007).

Moss, W. J. & Griffin, D. E. Global measles elimination. *Nat. Rev. Microbiol.* **4**, 900–908

(2006).

Mostafa, A., Abdelwhab, E. M., Mettenleiter, T. C. & Pleschka, S. Zoonotic Potential of Influenza A Viruses: A Comprehensive Overview. *Viruses* **10**, (2018).

Msaouel, P. *et al.* Clinical trials with oncolytic measles virus: Current status and future prospects. *Curr. Cancer Drug Targets* **18**, 177–187 (2018).

Mühlebach, M. D. Measles virus in cancer therapy. *Curr. Opin. Virol.* **41**, 85–97 (2020).

Nan, L. *et al.* On-Demand Droplet Collection for Capturing Single Cells. *Small* **16**, e1902889 (2020).

Nayak, D. P., Tobita, K., Janda, J. M., Davis, A. R. & De, B. K. Homologous interference mediated by defective interfering influenza virus derived from a temperature-sensitive mutant of influenza virus. *J. Virol.* **28**, 375–386 (1978).

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. *J. Virol.* **71**, 2138–2145 (1997).

Ohno, S., Ono, N., Takeda, M., Takeuchi, K. & Yanagi, Y. Dissection of measles virus V protein in relation to its ability to block alpha/beta interferon signal transduction. *J. Gen. Virol.* **85**, 2991–2999 (2004).

Oltean, H. N. *et al.* Sentinel Surveillance System Implementation and Evaluation for SARS-CoV-2 Genomic Data, Washington, USA, 2020-2021. *Emerg. Infect. Dis.* **29**, 242–251 (2023).

Ong, H. T. *et al.* Oncolytic measles virus targets high CD46 expression on multiple myeloma cells. *Exp. Hematol.* **34**, 713–720 (2006).

Ozer, E. A. *et al.* Multiple expansions of globally uncommon SARS-CoV-2 lineages in Nigeria. *Nat. Commun.* **13**, 688 (2022).

Panina, Y., Germond, A., David, B. G. & Watanabe, T. M. Pairwise efficiency: a new mathematical approach to qPCR data analysis increases the precision of the calibration curve assay. *BMC Bioinformatics* **20**, 295 (2019).

Pawelek, K. A. *et al.* Modeling within-host dynamics of influenza virus infection including immune responses. *PLoS Comput. Biol.* **8**, e1002588 (2012).

Peirson, S. N., Butler, J. N. & Foster, R. G. Experimental validation of novel and conventional approaches to quantitative real-time PCR data analysis. *Nucleic Acids Res.* **31**, e73–e73 (2003).

Petelski, A. A. *et al.* Multiplexed single-cell proteomics using SCoPE2. *Nat. Protoc.* **16**, 5398–5425 (2021).

Peterson, V. M. *et al.* Multiplexed quantification of proteins and transcripts in single cells. *Nat. Biotechnol.* **35**, 936–939 (2017).

Petrova, V. N. & Russell, C. A. The evolution of seasonal influenza viruses. *Nat. Rev. Microbiol.* **16**, 47–60 (2018).

Phipps, K. L. *et al.* Collective interactions augment influenza A virus replication in a host-dependent manner. *Nat Microbiol* **5**, 1158–1169 (2020).

Pinilla, L. T., Holder, B. P., Abed, Y., Boivin, G. & Beauchemin, C. A. A. The H275Y neuraminidase mutation of the pandemic A/H1N1 influenza virus lengthens the eclipse phase and reduces viral output of infected cells, potentially compromising fitness in ferrets. *J. Virol.* **86**, 10651–10660 (2012).

Piorkowski, G. *et al.* Development of generic Taqman PCR and RT-PCR assays for the detection of DNA and mRNA of β-actin-encoding sequences in a wide range of animal species. *J. Virol. Methods* **202**, 101–105 (2014).

Piorkowski, G. *et al.* Development of generic Taqman PCR and RT-PCR assays for the detection of DNA and mRNA of β-actin-encoding sequences in a wide range of animal species. *J. Virol. Methods* **202**, 101–105 (2014).

Poirier, E. Z. & Vignuzzi, M. Virus population dynamics during infection. *Curr. Opin. Virol.* **23**, 82–87 (2017).

Poon, L. L. M. *et al.* Quantifying influenza virus diversity and transmission in humans. *Nat. Genet.* **48**, 195–200 (2016).

Prakadan, S. M., Shalek, A. K. & Weitz, D. A. Scaling by shrinking: empowering single-cell 'omics' with microfluidic devices. *Nat. Rev. Genet.* **18**, 345–361 (2017).

Prakash, R. *et al.* Multiplex, Quantitative, Reverse Transcription PCR Detection of Influenza Viruses Using Droplet Microfluidic Technology. *Micromachines* **6**, 63–79 (2014).

Prakash, R. *et al.* Droplet Microfluidic Chip Based Nucleic Acid Amplification and Real-Time Detection of Influenza Viruses. *J. Electrochem. Soc.* **161**, B3083–B3093 (2014).

Rafael Sanjuán, P. D.-C. Genetic Diversity and Evolution of Viral Populations. *Encyclopedia of Virology* 53 (2021).

Raghwani, J., Thompson, R. N. & Koelle, K. Selection on non-antigenic gene segments of

seasonal influenza A virus and its impact on adaptive evolution. *Virus Evol* **3**, vex034 (2017).

Ramakers, C., Ruijter, J. M., Deprez, R. H. L. & Moorman, A. F. M. Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neurosci. Lett.* **339**, 62–66 (2003).

Ramos, I. *et al.* Innate Immune Response to Influenza Virus at Single-Cell Resolution in Human Epithelial Cells Revealed Paracrine Induction of Interferon Lambda 1. *J. Virol.* **93**, (2019).

Rato, S., Golumbeanu, M., Telenti, A. & Ciuffi, A. Exploring viral infection using single-cell sequencing. *Virus Res.* **239**, 55–68 (2017).

Rodpothong, P. & Auewarakul, P. Viral evolution and transmission effectiveness. *World J. Virol.* **1**, 131–134 (2012).

Rotem, A. *et al.* Evolution on the Biophysical Fitness Landscape of an RNA Virus. *Mol. Biol. Evol.* **35**, 2390–2400 (2018).

Ruijter, J. M. *et al.* Evaluation of qPCR curve analysis methods for reliable biomarker discovery: bias, resolution, precision, and implications. *Methods* **59**, 32–46 (2013).

Ruijter, J. M. *et al.* Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Res.* **37**, e45 (2009).

Russell, A. B., Kowalsky, J. R. & Bloom, J. D. Single-cell virus sequencing of influenza infections that trigger innate immunity. *bioRxiv* 437277 (2018) doi:10.1101/437277.

Russell, A. B., Elshina, E., Kowalsky, J. R., te Velthuis, A. J. W. & Bloom, J. D. Single-Cell Virus Sequencing of Influenza Infections That Trigger Innate Immunity. *J. Virol.* **93**, (2019).

Russell, A. B., Trapnell, C. & Bloom, J. D. Extreme heterogeneity of influenza virus infection in single cells. *Elife* **7**, (2018).

Russell, A. B., Elshina, E., Kowalsky, J. R., Te Velthuis, A. J. W. & Bloom, J. D. Single-Cell Virus Sequencing of Influenza Infections That Trigger Innate Immunity. *J. Virol.* **93**, (2019).

Russell, S. J. & Peng, K. W. Measles virus for cancer therapy. *Curr. Top. Microbiol. Immunol.* **330**, 213–241 (2009).

Russell, S. J. *et al.* Remission of disseminated cancer after systemic oncolytic virotherapy. *Mayo Clin. Proc.* **89**, 926–933 (2014).

Rutledge, R. G. & Stewart, D. A kinetic-based sigmoidal model for the polymerase chain reaction and its application to high-capacity absolute quantitative real-time PCR. *BMC*

Biotechnol. **8**, 47 (2008).

Rutledge, R. G. Sigmoidal curve-fitting redefines quantitative real-time PCR with the prospective of developing automated high-throughput applications. *Nucleic Acids Res.* **32**, e178 (2004).
Rybak-Wolf, A. *et al.* Neurodegeneration in human brain organoids infected with herpes simplex virus type 1. *bioRxiv* (2021) doi:10.1101/2021.03.05.434122.

Saikia, M. *et al.* Simultaneous multiplexed amplicon sequencing and transcriptome profiling in single cells. *Nat. Methods* **16**, 59–62 (2019).

Sambaturu, N. *et al.* Role of genetic heterogeneity in determining the epidemiological severity of H1N1 influenza. *PLoS Comput. Biol.* **14**, e1006069 (2018).

Sanjuán, R. The Social Life of Viruses. *Annu Rev Virol* (2021) doi:10.1146/annurev-virology-091919-071712.

Sanjuán, R. From molecular genetics to phylodynamics: evolutionary relevance of mutation rates across viruses. *PLoS Pathog.* **8**, e1002685 (2012).

Sanjuán, R. *et al.* Five challenges in the field of viral diversity and evolution. *Front. Virol.* **1**, 684949 (2021).

Sanjuán, R., Nebot, M. R., Chirico, N., Mansky, L. M. & Belshaw, R. Viral Mutation Rates. *J. Virol.* **84**, 9733–9748 (2010).

Santiago, G. A. *et al.* Genomic surveillance of SARS-CoV-2 in Puerto Rico enabled early detection and tracking of variants. *Commun. Med.* **2**, 100 (2022).

Scholtissek, C. Molecular evolution of influenza viruses. *Virus Genes* **11**, 209–215 (1995).

Schrag, S. J., Rota, P. A. & Bellini, W. J. Spontaneous mutation rate of measles virus: direct estimation based on mutations conferring monoclonal antibody resistance. *J. Virol.* **73**, 51–54 (1999).

Schulte, M. B. & Andino, R. Single-cell analysis uncovers extensive biological noise in poliovirus replication. *J. Virol.* **88**, 6205–6212 (2014).

Schwartz, S. L. & Lowen, A. C. Droplet digital PCR: A novel method for detection of influenza virus defective interfering particles. *J. Virol. Methods* **237**, 159–165 (2016).

Sen, A. *et al.* Innate immune response to homologous rotavirus infection in the small intestinal villous epithelium at single-cell resolution. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 20667–20672 (2012).

Shu, B. *et al.* Design and performance of the CDC real-time reverse transcriptase PCR swine flu panel for detection of 2009 A (H1N1) pandemic influenza virus. *J. Clin. Microbiol.* **49**, 2614–2619 (2011).

Sigal, D., Reid, J. N. S. & Wahl, L. M. Effects of Transmission Bottlenecks on the Diversity of Influenza A Virus. *Genetics* **210**, 1075–1088 (2018).

Song, Y. *et al.* Single cell transcriptomics: moving towards multi-omics. *Analyst* **144**, 3172–3189 (2019).

Spackman, E. *et al.* Development of a real-time reverse transcriptase PCR assay for type A influenza virus and the avian H5 and H7 hemagglutinin subtypes. *J. Clin. Microbiol.* **40**, 3256–3260 (2002).

Spiess, A.-N., Feig, C. & Ritz, C. Highly accurate sigmoidal fitting of real-time PCR data by introducing a parameter for asymmetry. *BMC Bioinformatics* **9**, 221 (2008).

Ståhl, P. L. *et al.* Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* **353**, 78–82 (2016).

Stern, A. *et al.* The Evolutionary Pathway to Virulence of an RNA Virus. *Cell* **169**, 35-46.e19 (2017).

Steuerman, Y. *et al.* Dissection of Influenza Infection In Vivo by Single-Cell RNA Sequencing. *Cell Syst* **6**, 679-691.e4 (2018).

Stoeckius, M. *et al.* Simultaneous epitope and transcriptome measurement in single cells. *Nat. Methods* **14**, 865–868 (2017).

Stray, S. J. & Air, G. M. Apoptosis by influenza viruses correlates with efficiency of viral mRNA synthesis. *Virus Res.* **77**, 3–17 (2001).

Sun, J. *et al.* Single cell heterogeneity in influenza A virus gene expression shapes the innate antiviral response to infection. *PLoS Pathog.* **16**, e1008671 (2020).

Sun, J. *et al.* Single cell heterogeneity in influenza A virus gene expression shapes the innate antiviral response to infection. *PLoS Pathog.* **16**, e1008671 (2020).

Tao, H., Li, L., White, M. C., Steel, J. & Lowen, A. C. Influenza A Virus Coinfection through Transmission Can Support High Levels of Reassortment. *J. Virol.* **89**, 8453–8461 (2015).

Tao, Y. *et al.* Artifact-Free Quantification and Sequencing of Rare Recombinant Viruses by Using Drop-Based Microfluidics. *Chembiochem* **16**, 2167–2171 (2015).

Tao, Y. *et al.* Rapid, targeted and culture-free viral infectivity assay in drop-based microfluidics. *Lab Chip* **15**, 3934–3940 (2015).

Tao, Y. *et al.* Rapid, targeted and culture-free viral infectivity assay in drop-based microfluidics. *Lab Chip* **15**, 3934–3940 (2015).

Taubenberger, J. K. & Kash, J. C. Influenza virus evolution, host adaptation, and pandemic formation. *Cell Host Microbe* **7**, 440–451 (2010).

Tellinghuisen, J. & Spiess, A.-N. qPCR data analysis: Better results through iconoclasm. *Biomol Detect Quantif* **17**, 100084 (2019).

Theberge, A. B. *et al.* Microdroplets in microfluidics: an evolving platform for discoveries in chemistry and biology. *Angew. Chem. Int. Ed Engl.* **49**, 5846–5868 (2010).

Timm, A. & Yin, J. Kinetics of virus production from single cells. *Virology* **424**, 11–17 (2012).

Timm, A. C., Warrick, J. W. & Yin, J. Quantitative profiling of innate immune activation by viral infection in single cells. *Integr. Biol.* **9**, 782–791 (2017).

Truong, T. & Kelly, R. T. What's new in single-cell proteomics. *Curr. Opin. Biotechnol.* **86**, 103077 (2024).

Varble, A. *et al.* Influenza A virus transmission bottlenecks are defined by infection route and recipient host. *Cell Host Microbe* **16**, 691–700 (2014).

Vignuzzi, M. & López, C. B. Defective viral genomes are key drivers of the virus–host interaction. *Nature microbiology* (2019).

Vignuzzi, M. & López, C. B. Defective viral genomes are key drivers of the virus–host interaction. *Nature Microbiology* **4**, 1075–1087 (2019).

Vignuzzi, M., Stone, J. K., Arnold, J. J., Cameron, C. E. & Andino, R. Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. *Nature* **439**, 344–348 (2006).

von Magnus, P. Incomplete Forms of Influenza Virus. in *Advances in Virus Research* (eds. Smith, K. M. & Lauffer, M. A.) vol. 2 59–79 (Academic Press, 1954).

Wang, C. *et al.* Cell-to-Cell Variation in Defective Virus Expression and Effects on Host Responses during Influenza Virus Infection. *MBio* **11**, (2020).

Wang, H., Xin, X., Zheng, C. & Shen, C. Single-Cell Analysis of Foot-and-Mouth Disease Virus. *Front. Microbiol.* **11**, 361 (2020).

Wang, Y. *et al.* Directed evolution: Methodologies and applications. *Chem. Rev.* **121**, 12384–12444 (2021).

Warrick, J. W., Timm, A., Swick, A. & Yin, J. Tools for Single-Cell Kinetic Analysis of Virus-Host Interactions. *PLoS One* **11**, e0145081 (2016).
Wilkinson, E. *et al.* A year of genomic surveillance reveals how the SARS-CoV-2 pandemic unfolded in Africa. *Science* **374**, 423–431 (2021).

Xin, X. *et al.* Single-Cell Analysis of the Impact of Host Cell Heterogeneity on Infection with Foot-and-Mouth Disease Virus. *J. Virol.* **92**, (2018).

Xu, X. *et al.* Microfluidic Single-Cell Omics Analysis. *Small* **16**, e1903905 (2020).

Xue, K. S., Moncla, L. H., Bedford, T. & Bloom, J. D. Within-Host Evolution of Human Influenza Virus. *Trends Microbiol.* **26**, 781–793 (2018).

Zanini, F., Pu, S.-Y., Bekerman, E., Einav, S. & Quake, S. R. Single-cell transcriptional dynamics of flavivirus infection. *Elife* **7**, (2018).

Zath, G. K and Thomas, M.M *et al.* Influenza A viral burst size from thousands of infected single cells using droplet quantitative PCR (dqPCR). *PLoS Pathog.* **20**, e1012257 (2024).

Zhang, H. *et al.* Isolation and Analysis of Rare Norovirus Recombinants from Coinfected Mice Using Drop-Based Microfluidics. *J. Virol.* **89**, 7722–7734 (2015).

Zhou, W.-M. *et al.* Microfluidics applications for high-throughput single cell sequencing. *J. Nanobiotechnology* **19**, 312 (2021).

Zhu, H., Fohlerová, Z., Pekárek, J., Basova, E. & Neužil, P. Recent advances in lab-on-a-chip technologies for viral diagnosis. *Biosens. Bioelectron.* **153**, 112041 (2020).

Zhu, Y. *et al.* Nanodroplet processing platform for deep and quantitative proteome profiling of 10-100 mammalian cells. *Nat. Commun.* **9**, 882 (2018).

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

Zilionis, R. *et al.* Single-cell barcoding and sequencing using droplet microfluidics. *Nat. Protoc.* **12**, 44–73 (2017).