Title
Targeted Virus Detection and Enrichment Using Droplet Microfluidics

Author
Thompson, Catharine Shea

Publication Date
2016

Peer reviewed|Thesis/dissertation

Targeted Virus Detection and Enrichment Using Droplet Microfluidics

by

Catharine Shea Thompson

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Bioengineering

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

AND

UNIVERSITY OF CALIFORNIA, BERKELEY

Approved:

[redacted]

Acknowledgements

Elements of this dissertation have been published elsewhere or are in preparation for submission to a peer-reviewed journal. Chapter 2 was published in *Journal of Physics D: Applied Physics*. Chapter 3 was published in *Lab on a Chip*. Chapter 4 and Chapter 5 have been submitted to peer-reviewed journals, and Chapter 6 is in preparation for submission to a peer-reviewed journal for publication.

First of all, I want to acknowledge my advisor, Dr. Adam Abate, for his years of mentorship and for giving me the opportunity to conduct this research. Adam is a brilliant scientist and researcher who inspires his students. He fosters collaborations with other research groups, which creates new opportunities to apply the technologies we create in the laboratory to address open challenges in a range of life science fields.

I want to thank the members of the Abate lab for always being available and willing to help by giving advice, assistance, or precious surfactant. In particular, I would like to thank Dr. David Sukovich, a Staff Research Associate in the lab. It was a privilege to work alongside David, and his mentorship on issues both scientific and personal was invaluable. I also want to thank my fellow graduate students Shaun Lim, Freeman Lan, John Haliburton, and Tuan Tran. As the lab grew and changed through the years, they were steadfast companions.

I want to thank Professor Dorian Liepmann and Dr. Joseph Wong for serving on my dissertation committee. Dorian provided guidance throughout my graduate career, from interview weekend to dissertation committee. Joe Wong was also my valued collaborator on the HIV-PACS project, which brought my research closest to the bedside and most inspired me during graduate school. I would also like to thank Professors Dorian Liepmann, Joe DeRisi, and Shuvo Roy for serving on my qualifying exam committee and providing valuable research guidance.

I would like to thank my family for their love, humor, and encouragement throughout my years as a graduate student. This long and difficult journey was only possible with them in my corner. I especially want to thank my parents for nurturing my ambitions and allowing me to carve my own path.

Lastly, and most importantly, I want to thank my husband Kevin. His jokes lift my spirits, his advice helps me see a way forward, and his love comforts me always.

Targeted Virus Detection and Enrichment Using Droplet Microfluidics

Abstract

Viruses impact every form of life on earth, from causing disease in individuals to influencing to applying evolutionary stresses and influencing environmental communities. There are an incredible number of viruses in the biosphere that virological research seeks to characterize in order to understand their health and environmental impacts. Next-generation sequencing (NGS) is a powerful tool for characterizing viral genomes, yet it is imperative to pre-enrich complex samples for viruses of interest in order to optimize coverage. This dissertation describes the development of several new technologies for targeted quantification and enrichment of viruses using droplet microfluidics. These technologies represent a general platform for viral enrichment that can be applied directly or used upstream of other technologies such as NGS. We describe methods for increasingly complex, targeted enrichment, starting with the enrichment of specific DNA molecules and specific viruses. We then describe a viral enrichment strategy capable of identifying microbial viral hosts. Last, we detail a method for enrichment of mammalian cells infected with viruses of interest and describe how this technology can be used to quantify and characterize the latent HIV reservoir. We anticipate that these techniques will enable exploration of numerous virological avenues including viral diversity, viral evolution, host-virus relationships, and viral co-infections.

Table of Contents

List of Tables

List of Figures

List of Figures (cont'd)

List of Figures (cont'd)

Chapter 1: Introduction

Viruses are incredibly diverse organisms that impact all forms of life on Earth. They are ubiquitous in that they influence nearly all aspects of our lives by shaping the environment, our immune responses, and even our genomes. However, viruses are challenging to study due to their small size and mass, precluding almost all direct imaging or molecular analysis of individual virions. Moreover, the overwhelming majority of viruses cannot be cultured, which impedes the isolation, replication, and study of interesting new species.

Next-generation sequencing (NGS) is a powerful tool for charactering viruses at the genomic level. It is crucial to enrich for the viral species of interest before performing NGS analysis because other abundant DNA sequences in the sample, such as genomic fragments of cellular DNA, are often larger than viral genomes and dominate the sequencing data. Traditional enrichment methods for viruses rely on cell culture, yet fewer than 1% of viruses are cultivable, which limits virological study to a subset of species. A common method for targeted enrichment is hybridization-based capture, but this technique is less effective at recovering large genomic regions and requires the design of a tiled oligonucleotide library, so this method cannot be used for exploratory studies of new viruses with unknown sequences. Another targeted enrichment technique is Fluorescence In Situ Hybridization coupled with Fluorescence-Activated Cell Sorting (FISH-FACS). While FISH-FACS does not rely on viral cultivation, it requires sample fixation, which can fragment DNA and introduce mutations during PCR, affecting the quality of downstream NGS data. The advancement of virological research through NGS requires the development of new targeted viral enrichment technologies that preserve genomic integrity and do not require cultivation.

Droplet microfluidics is a growing field in microfluidics that isolates cells or biomolecules into picoliter volume droplets surrounded in an inert carrier oil. As such, the droplets are analogous to conventional test tubes for reactions, except their smaller volumes afford order-of-magnitude improvements in throughput, cost, and sensitivity. Droplet microfluidics opens a new frontier in biological research, enabling a wide variety of experiments such as directed evolution, drug screening, and single cell analysis. We propose to combine the ultrahigh-throughput nature of microfluidics with the specificity of PCR reactions to create a generalizable platform for targeted viral enrichment.

This dissertation details multiple demonstrations of the targeted quantification and enrichment of viruses using droplet microfluidics and PCR. Several different methods are described, but all of the applications are based on the general concept illustrated in Figure 1.1. First, the viral sample is partitioned into droplets such that the viruses are individually encapsulated. The encapsulated viruses are then subjected to PCR with primers that are specific to the desired viral species. The PCR reaction produces a detectable fluorescent signal in droplets containing the virus of interest, through either TaqMan probes or DNA stain. The fluorescent signal is detected and used to quantify the amount of virus in the sample. The signal also triggers the sorting of droplets containing the virus of interest, which results in an enriched viral sample that is available for additional downstream analyses including qPCR and next-generation sequencing. We anticipate that these techniques will enable exploration into a variety of virological concerns including viral diversity, viral evolution, host-virus relationships, and viral co-infections.

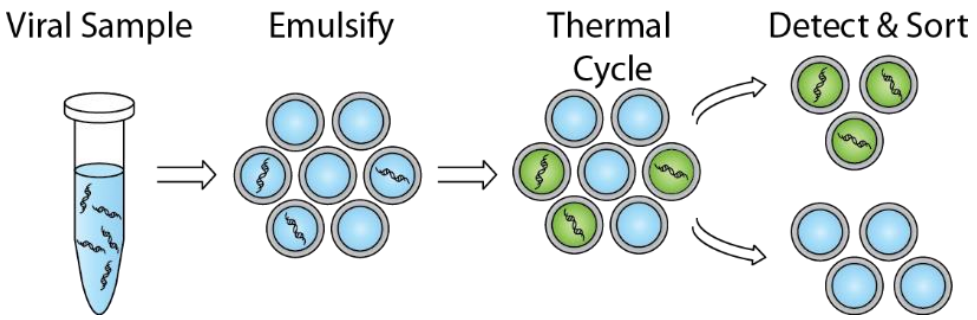


Figure 1.1 Generalized workflow for targeted virus quantification and enrichment

In the chapters that follow, we first present background on droplet microfluidics, then describe development of new microfluidic enrichment methods, and finally detail several applications of targeted quantification and enrichment of viruses.

Chapter 2 provides background on droplet microfluidics and its potential application to a variety of biological problems. We first describe the process of device fabrication, droplet generation, and cell encapsulation. We then detail various microfluidic modules for manipulating droplets, such as adding reagent and sorting the droplets. Last we present biological applications that droplet microfluidics is particularly well suited to address: namely, applications that require ultrahigh throughput such as screening massive libraries or analyzing millions of cells.

Chapter 3 describes an innovative fabrication method for bonding microfluidic devices using adhesive tapes and baking. This new technique is convenient for rapid prototyping and uses inexpensive and widely available materials, in contrast to other bonding methods that require specialized equipment.

Chapter 4 describes a method for targeted enrichment of DNA molecules called 3DPCR. This technique simultaneously quantifies and sorts DNA molecules of interest and is compatible with TaqMan probes for multiplexed detection. We show the enrichment of one DNA sequence of interest from a heterogeneous sample and confirm the enrichment using qPCR.

Chapter 5 applies the method described in Chapter 4 to quantify and enrich for viruses of interest. We demonstrate how this technique, PCR-Activated Virus Sorting (PAVS), is capable of specific enrichment of one viral species and describe how PAVS can be used to isolate single virions for downstream analysis.

Chapter 6 describes a workflow for identification of viral hosts and studying virus-host interactions, as it is unclear what hosts the vast majority of viruses infect. We demonstrate the identification of a viral host species from a heterogeneous sample and quantify the enrichment factor.

Chapter 7 describes PCR-Activated Cell Sorting (PACS). Rather than detecting free viruses or viruses within microbial hosts, the PACS workflow described in this chapter can detect viruses within mammalian cells. This workflow is necessary for detecting and enriching many viruses that are greatly important for human heath such as Human Immunodeficiency Virus (HIV). In addition to enriching for the viral genomes, the PACS workflow implemented in this manner also produces enriched infected cell genomes. In this chapter, we first describe the PACS workflow. Next we detail several optimization strategies that improve the robustness of the PACS workflow. Last we describe the potential application of PACS for the quantification and enrichment of HIV-infected T-cells.

Chapter 2: From tubes to drops: Droplet-based microfluidics for ultrahigh-throughput biology

The following section is reprinted from "From tubes to drops: Droplet-based microfluidics for ultrahigh-throughput biology" by Tuan M. Tran*, Freeman Lan*, C. Shea Thompson*, and Adam R. Abate. The article was published in Volume 46, Number 11 of the Journal of Physics D: Applied Physics on 22 February 2013. This review paper details the background of droplet-based microfluidics and its potential application to a variety of biological problems. Tuan M. Tran, Freeman Lan, C. Shea Thompson researched the relevant literature and co-wrote the publication. Adam Abate provided guidance in outlining the scope of the review and supervised the project.

2.1 Abstract

Droplet-based microfluidics holds enormous potential for transforming the way that many biological screens are performed by affording unprecedented increases in screening throughput and reductions in reagent usage. In this Review, we describe this maturing field and the pioneering work that has laid the foundation for its application to ultrahigh-throughput biological analysis. We begin by introducing the basic elements of the approach and describe the numerous microfluidic components that have been developed for droplet manipulation, with special emphasis on the ones most useful for ultrahigh-throughput analysis. We conclude with a discussion of the first demonstrations of this approach to perform novel, ultrahigh-throughput biological screens.

**Authors contributed equally*

2.2 Introduction

There are many examples in biological research in which the major barrier to progress is the need to screen massive numbers of separate biological reactions. For example, in drug discovery applications it is often necessary to screen hundreds-of-thousands of compounds, each of which must be separately tested for activities of interest[1–3]. When engineering proteins through directed evolution, it is often necessary to screen millions of variants of the target protein to find the rare variants with the best activity[4–7]. To detect rare cells or pathogens in a mixed sample, it is often necessary to screen and classify hundreds-of-millions of candidate cells[8]. In all of these examples, and many others like them, the ability to perform ultrahigh-throughput screening on a system-wide scale is essential.

There is now a growing arsenal of tools available to researchers in the biological sciences that are amenable to ultrahigh-throughput, system-scale analysis. To analyze systems at the genomic and transcriptional level, Next Generation Sequencing (NGS) has been transformative because it is able to sequence hundreds-of-millions of DNA molecules, allowing every variant in the system to be characterized. For proteomic analysis, mass spectrometry is becoming an ever more powerful tool, allowing the detection and classification of hundreds of proteins in a complex mixture, and even the determination of protein-protein interactions[9]. This has opened up entirely new frontiers in systems biology, in which complex protein interaction networks are systematically constructed from the rapidly elucidated interactions among pairs of proteins. To analyze systems at the cellular level, fluorescence-activated cell sorting (FACS) is a universal tool, allowing millions of single cells to be individually screened and sorted according to their fluorescence and light-scattering properties.

There is, nevertheless, an important and unmet need in ultrahigh-throughput biological studies, which is the ability to rapidly execute and screen liquid-phase reactions. For example, when enhancing the activity of an enzyme through directed evolution, a critical step is characterizing the catalytic activity of each enzyme variant in the pool. However, because the product of an enzymatic reaction is a molecule that is released into the surrounding solution, there is no physical linkage between the enzyme, its gene, and the quantity of product it produces, precluding the direct use of technologies like FACS for the ultrahigh-throughput screening. Instead, the reactions must be modified to link gene, protein, and product[6,10–13], an extremely challenging task; alternatively, each variant can be tested in a separate well using plate-based screening[5,7,14–20]. Well-plate formats, however, due to their limited throughput, can only screen ~105 variants in total; this is far too small to cover the sequence space of even the active site of most enzymes, making it ineffective for many directed evolution studies. As an alternative to well-plate formats, electrowetting-on-dielectric (EWOD) is a growing field in microfluidics that compartmentalizes reactions in droplets positioned on a checkerboard-like array. Using electrowetting forces, the droplets can be moved around the board and split, merged, and passed over sensors to probe their contents. The precision, automation, and flexibility of this approach hold immense potential for liquid reaction screening, but current technologies are still limited in throughput to just a few reactions per second[21]. Increasing the throughput of liquid-reaction processing would benefit a number of important biological screens, including drug discovery, genomics, and the development of therapeutic antibodies, to name just a few examples.

In this Review, we describe the burgeoning field of microchannel droplet-based microfluidics, a new frontier in microfluidics that holds potential for transforming the way that many biological screens are conducted. In this approach, microdroplets, tiny spheres of aqueous

liquid dispersed in an inert carrier oil, are used as "test tubes" for reactions with single cells and biomolecules. In many ways, these devices are similar to automated well-plate screening platforms, except that the wells are thousands of times smaller and processed at rates thousands of times faster. The combination of small volumes and massive screening throughput allows screening on a scale that is infeasible with conventional approaches – in which millions of reactions are screened in hours using microlitres of total reagent by flowing them at high velocities through precision-fabricated microfluidic channels. Here, we will describe the basic elements of the approach and the multitude of microfluidic components that have been developed for droplet manipulation, screening, and sorting, with special emphasis on the components most amenable to ultrahigh-throughput applications: ones in which picoliter-sized droplets are screened at kilohertz rates. We conclude with a discussion of the recent efforts of researchers to build integrated platforms and their application to biological screens that are the first of their kind.

2.3 Fabricating droplet-based microfluidic devices

Photolithography of SU-8 masters

A key explanation for the explosion of microfluidics in the last decade is the development of soft lithography in poly(dimethylsiloxane) (PDMS), a fabrication process that allows creation of microfluidic devices with a range of channel geometries and with precision and ease. PDMS has several properties that make it attractive as a material for microfluidic devices: it is naturally hydrophobic, optically transparent, minimally fluorescent at UV and visible wavelengths, and chemically inert[22]. It is, however, also permeable to vapors and gases, which can be problematic for certain applications. Alternative materials that have been used for creating droplet microfluidic devices for biological applications include glass[23] and PMMA[24]. We will focus on devices

fabricated in PDMS, because they are by far the most widely used and the best for ultrahigh-throughput droplet-based screening in academic research settings.

PDMS devices are normally fabricated by molding them from masters containing positive (protruding) channels of the epoxy SU-8 glued to a silicon wafer[25]. The masters are fabricated through a lithographic process in which the photocurable epoxy is spin-coated onto the silicon wafer at a controlled thickness. The wafer is then baked to drive excess solvent out of the epoxy and to cause it to harden, but not crosslink, when cooled. The epoxy-coated wafer is then covered with a "mask" consisting of a printout of the desired microfluidic device in inverse – that is, in which most of the mask is coated with UV-absorbent ink, but the portions that are to become the positive channels are transparent. The wafers are then exposed to collimated UV light, such that the light passes through the transparent regions but is blocked by the absorbent regions; this crosslinks the SU-8 under the transparent regions. The wafer can then be "developed" by bathing it in a solvent that dissolves uncrosslinked SU-8, leaving behind positive features in the shape needed to mold the PDMS device[26]. This approach produces planar microfluidic channels with rectangular cross sections; however, by iterating coatings of different thickness with exposures through different masks, it is possible to fabricate non-planar devices in which the channel heights vary too[27]. In the majority of research labs utilizing this process, clean room facilities are used, although they are not absolutely required: Depending on the sizes of features in the device and the particular use for which it is intended, often a tidy research lab suffices.

Molding PDMS replicates

Once the master has been fabricated, the PDMS mold can be replicated. This is accomplished by preparing a batch of silicone elastomer (Sylgard 184) and pouring it over the

master. The mold is evacuated to remove entrained air bubbles and baked. The baking accelerates crosslinking of the PDMS elastomer, causing it to solidify and become a transparent rubber. It is then sliced and peeled from the master, punched with inlet ports, and washed and bonded to a solid support, such as a glass slide.

Bonding and sealing PDMS channels

The bonding of the PDMS device to a solid support is a critical step in the fabrication process because it adds the final wall that encloses the channels. The most common method for bonding PDMS channels for ultrahigh-throughput biological applications is oxygen plasma treatment. In this approach, a fully-cured PDMS replicate is treated with oxygen plasma immediately before bonding it to another piece of PDMS or glass. The oxygen plasma makes the surfaces of the PDMS reactive so that when they are placed into contact they irreversibly bond. Plasma bonding produces strong bonds in a matter of minutes. However, it also requires an oxygen plasma cleaner, an expensive piece of hardware that is not available in most research labs. A less-expensive alternative is a handheld corona wand[28], which bonds devices through a similar process at a fraction of the cost. However, because the treatment is performed by hand in the atmosphere rather than in a controlled oxygen environment, the results are less consistent. Both methods make the channels temporarily hydrophilic[29]. To enable the formation of the aqueous-in-oil emulsions that are used in ultrahigh-throughput biological applications, the channels must be made hydrophobic, which is normally achieved using chemical treatments.

Channel wettability

To form droplets in lithographically-fabricated microfluidic channels, the wetting properties of the channels are critical. In planar microfluidic devices like the ones normally fabricated in PDMS, all fluids are initially in contact with the channel walls. To form aqueous droplets, the channel walls must therefore be hydrophobic to allow the oil phase to lift the aqueous phase from the walls, surround it, and encapsulate it into drops. Hydrophobic channels can be obtained by plasma bonding the PDMS channels to a PDMS surface and baking the device for several days at 65°C, during which time the channels revert to their native hydrophobic state[29]. Alternatively, to make the channels hydrophobic more quickly, they can be functionalized with hydrophobic silanes or the glass treatment Aquapel.

Oils and surfactants

Just as important as the aqueous droplets that comprise the "test tubes" is the oil phase surrounding the droplets and comprising the "walls" of the test tubes. The carrier phase must, ideally, allow the droplets to be stable against coalescence, have a viscosity that is close to that of water, and be inert with respect to the biological reagents contained in the drops and the material of which the device is composed. Several oils have been used in droplet-based microfluidics, each with their own pros and cons.

Low viscosity silicone oils swell PDMS[30,31], changing the cross-sectional dimensions of the channels and influencing the flow properties of the devices[32]. This also depletes the carrier phase from the channels, limiting on-chip incubation time and interfering with droplet recovery[33]. Silicone oils, however, can be used in microfluidic devices fabricated in glass[23,34], which are impermeable to these oils; however, glass devices are much harder to fabricate than PDMS devices and, thus, less widely used. High viscosity silicone oils can be used in PDMS devices with minimal

swelling at the expense of significantly increasing the pressures required to pump them through the microchannels. Hydrocarbon oils can also be obtained in a range of viscosities and have the benefit that there are a large number of commercially-available surfactants for them that can stabilize aqueous-in-oil emulsions. However, they also swell PDMS[32,33] and tend to exhibit poor retention of encapsulated organic reagents, which are often partially soluble in these oils[35].

By far the preferred oils for biological applications of droplet-based microfluidics are fluorocarbon oils, because even low viscosity versions of these oils do not swell PDMS[32]. In addition, they tend to exhibit excellent retention of reagents in the drops[36,37] and have high solubility for gases, allowing oxygen and carbon dioxide to passively diffuse in and out of the drops, for unperturbed cellular respiration[38–40]. This allows yeast (Figure 2.1A)[39], algae[41], mammalian cells[40,42], and even the multicellular organism C. elegans[42] to survive in fluorocarbon oil emulsions for hours after encapsulation. A disadvantage of fluorocarbon oils, however, is that, due to their much lower prevalence compared to silicone and hydrocarbon oils, there are few commercially-available surfactants for stabilizing aqueous-in-fluorocarbon emulsions. Surfactants are essential for reducing the surface tension of the oil-water interface[43] and minimizing droplet coalescence[33]. The choice of which surfactant to use is also imperative for limiting the transfer of reagents between drops[44,45]. A comprehensive review of surfactants for droplet-based microfluidics is available[46]. The surfactants utilized in droplet-based microfluidics normally consist of a hydrophilic head group and hydrophobic tail. The amphiphilic character of these molecules allows them to assemble at the oil-water interface of the droplet, thereby lowering its interfacial tension and enhancing stability[43,47] as depicted in Figure 2.1B. The chemical properties of the head group of the surfactant impact the biocompatibility of the droplet interface. Surfactants with non-ionic head groups, for instance, have been found to minimize the adsorption of

macromolecules, like proteins and DNA, to the droplet interface, minimally impacting biological assays performed in the drops[47,48]. Several fluorosurfactants that can be readily synthesized in the lab have been described, as has their effectiveness at stabilizing emulsions and yielding biocompatible droplets[42,43,47–49]. Additives to the aqueous phase can also enhance biocompatibility by increasing the retention of small molecules in the droplets and minimizing adsorption at the oil-water interface[40,50,51]. The choice of the surfactant should be made with the oil that is to be used since the properties of the resultant emulsion depend on the combination[24,52]. Different oils can be mixed to optimize the properties of the emulsion for the particular application[53] and methods have been described for easily characterizing the properties of the combination that has been selected[52].

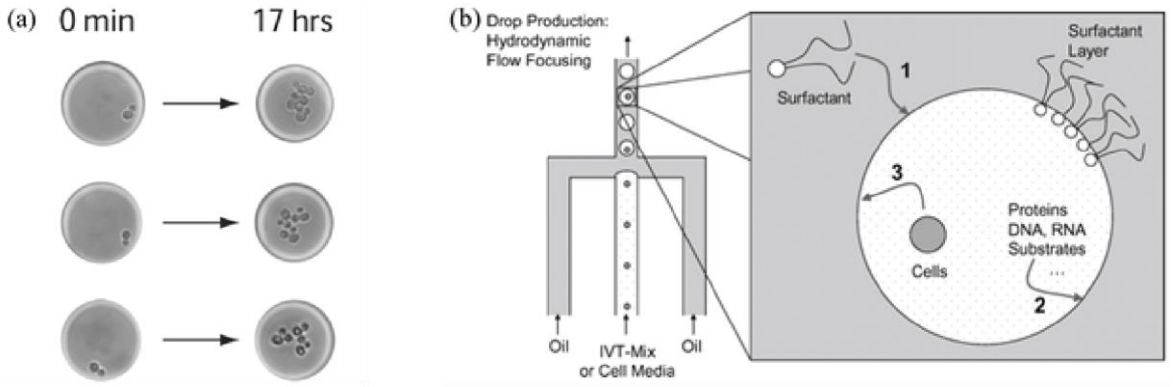


Figure 2.1 Fluorinated oils and non-ionic surfactants are currently thought to afford the best combination of properties for biological applications of droplet-based microfluidics, and are thus the most commonly used combination. They have been demonstrated to be compatible with different kinds of cells, including mammalian cells, bacteria, and yeast cells, which can still divide and proliferate long after encapsulation, (A). Adapted from [47] with permission of The Royal Society of Chemistry. The surfactants stabilize the droplets by adsorbing to the oil-water interface, lowering its interfacial tension and coating the inner surface with a hydrophilic, non-ionic, and

biocompatible layer, such as polyethylene glycol, that is resistant to protein adsorption and non-toxic to cells, (b). Reproduced from [47] with permission of The Royal Society of Chemistry.

2.4 Generating droplets and encapsulating cells and biomolecules

Droplet generation geometries

The starting point for most droplet-based microfluidic screens is droplet generation[54–56]. There are three common droplet generation geometries in microfluidics: co-flow, T-junction, and flow focus drop formation, each illustrated in Figure 2.2. In co-flow drop formation, the dispersed phase is injected through a small capillary centered within a larger capillary, flowing parallel to the flow of the continuous phase, as shown in Figure 2.2A[57,58]. Droplets are generated by the viscous shear of the continuous phase over the dispersed phase in a process that resembles a dripping faucet; as the emerging droplet grows, the viscous drag of the continuous phase increases. This continues until the drag is equal to the interfacial tension force adhering the base of the droplet to the capillary tip, at which point a droplet buds off and is carried downstream[59]. In co-flow drop formation, the flows are "unconfined" in the sense that the outer capillary is much larger than the inner capillary and the droplets that are formed; consequently, the drop formation mechanism depends mainly on viscous shear and surface tension, and interactions with the outer capillary wall can be neglected[60]. A disadvantage of this geometry is that it is difficult to fabricate with lithographic processes because it requires the inner capillary to be smaller than, and nested within, the outer capillary. As a result, this geometry is rarely used in ultrahigh-throughput biological applications, which typically utilize microfluidic devices fabricated with soft lithography and, thus, consisting of planar channel networks.

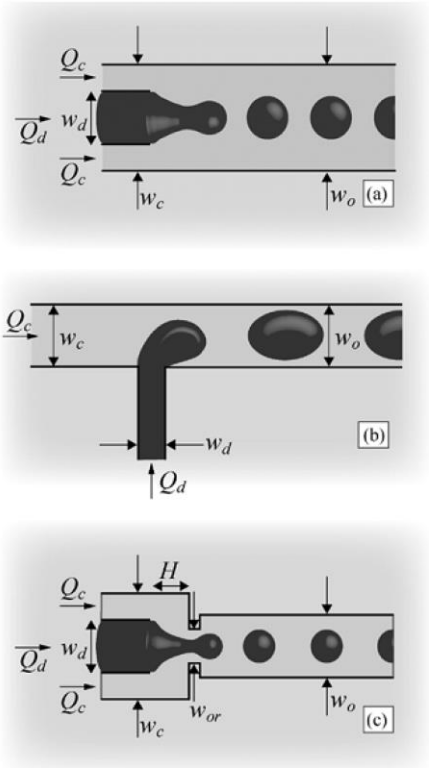


Figure 2.2 The three most commonly used droplet-formation geometries in microfluidics are co-flow (A), T-junction (B), flow-focus (C) drop formation. T-junction and flow focus drop formation are the most widely used because they are easily fabricated in PDMS devices consisting of planar microchannels. Important dimensions for determining droplet size are labeled including w_d, w_c, w_o, which are the widths of the dispersed, continuous and outer phase streams, respectively. w_{or} is the width of the orifice. Q_d and Q_c are the dispersed and continuous phase flow rates. Reproduced with permission from 54.

A more commonly used drop formation geometry, and one that is an example of "confined" drop formation, is the T-junction. In this geometry, the dispersed phase is injected from a channel that is perpendicular to the channel carrying the continuous phase, as illustrated in Figure 2.2B[61]. The mechanism by which monodisperse drops are formed depends on the Capillary number (Ca) of the flow. At high Ca, the large viscous drag of the continuous phase shears droplets of the

dispersed phase from the inlet channel. A characteristic of drops formed by this mechanism is that they are smaller than the channel. Alternatively, at low Ca drops can also be produced through a mechanism of plugging and squeezing[62]. In this mechanism, the emerging tip of the dispersed phase blocks the downstream channel to the flow of the continuous phase; this causes the pressure to rise in the continuous phase which, in turn, squeezes on the dispersed phase and pinches off a drop. In this mechanism pressure fluctuations are generated in the continuous and dispersed phases and the drops that are formed are always larger than the downstream channel. For the majority of cases in which T-junctions are used, plugging and squeezing is the primary mechanism of drop formation.

Another drop formation geometry is the flow focus device[63]. In this geometry, the dispersed phase is introduced from one channel and the continuous phase from channels on either side. The fluids are focused through an "orifice," where droplets are formed. There are several variations on this geometry, including one in which a small constriction is added at the orifice, as pictured in Figure 2.2C, and in which the constriction is omitted, yielding a straight "throat." Both variations tend to form droplets with comparable properties for the majority of flow rates. Like the T-junction, the mechanism of drop formation in the flow focus geometry depends on the Ca of the flow: At low Ca the drops form through a process of plugging and squeezing while at high Ca, they form through a process that depends primarily on viscous drag and interfacial tension[64].

At very high flow rates all three drop makers exhibit "jetting," in which a long tube of the dispersed phase jets through the drop maker and breaks up downstream due to the Rayleigh-Plateau instability[59,65,66], yielding somewhat polydisperse emulsions. The transition to jetting thus sets an upper limit to the rate at which monodisperse droplets can be formed. While this rate can exceed kilohertz for these drop makers, in some instances it is desirable to form drops even faster

and, thus, several strategies have been invented to increase drop rates. One strategy is to form large drops at kilohertz rates that are then split into small drops using geometrically-mediated breakup[67]. This increases the droplet production rate by a factor equal to the number of daughter drops formed from each large drop, and has been used to increase net throughput significantly[68]. Another strategy is to exploit the periodic formation of air bubbles to trigger the periodic breakup of an otherwise stable jet[69]. This allows greatly increased drop formation rates because air bubbles, due to their low viscosity and mass, can be formed at rates much faster than droplets of most liquids and, thus, can trigger the breakup of the jet at rates much faster than it would break up spontaneously. To increase the production rate further, drop makers can also be parallelized. Rather than operating one drop maker at a time, several can be operated simultaneously to increase the drop production rate by many times[23,70,71]. For a thorough review of rapid emulsion techniques, see the following reference[55]. Other useful techniques are valve-based flow focusing, which allows drop size and frequency to be adjusted without changing flow rates[72], and step emulsification, which allows the formation of very small droplets[73,74].

Encapsulating cells and biomolecules

Perhaps the greatest impact of droplet-based microfluidic techniques will be in their ability to perform massively-parallel analysis on populations of single cells. A critical step in this analysis is achieving controlled, efficient encapsulation of the cells in microdroplets. The simplest and most common way to achieve this is limiting dilution, in which a dilute suspension of cells is emulsified such that several drops are formed for every cell in the solution. Under such conditions, the cells are loaded randomly in a process governed by Poisson statistics[75]. This allows one to guarantee that an acceptably small fraction of the droplets contain more than a single cell, at the cost that the

majority are empty and unusable with only a small fraction containing single cells. This tradeoff, however, is often agreeable due to the simplicity of the method, its robustness in operation, and the enormous quantity with which droplets can be formed[76]. When this inefficiency is unacceptable, other methods can be implemented to greatly increase single-cell encapsulation efficiency.

One strategy for increasing encapsulation efficiency is to organize the cells prior to encapsulating them in the drops. This can be accomplished using inertial microfluidics to order the cells into a periodically spaced line that travels into the droplet maker at a constant velocity; by matching the periodicity of the cells to that of the drop formation, it is possible to encapsulate a controlled number of cells in every drop[76,77], as illustrated in Figure 2.3. The ordering of the cells is achieved passively by flowing the cells at high velocity through a long, narrow channel, making it suitable for ultrahigh-throughput applications. However, the high flow rates make this strategy difficult to integrate into complex microfluidic networks, which often contain components that require much lower flow rates. Another challenge is that in many biological assays the amount of reagent available is limited, making it difficult to achieve steady encapsulation before the regents are exhausted.

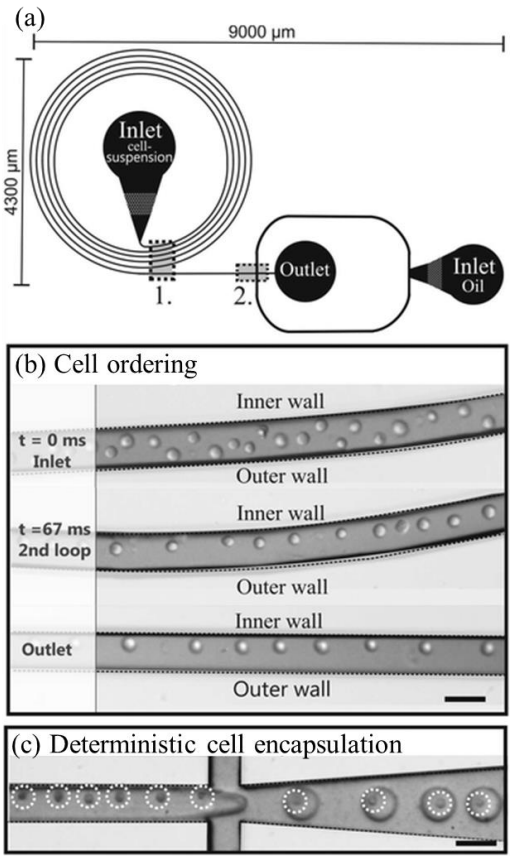


Figure 2.3 Cell encapsulation in droplet-based microfluidics is most often achieved using limiting dilution, so that some drops contain single cells but most are empty. Newer methods are allowing much greater encapsulation efficiency by exploiting inertial microfluidics (A). A spiral geometry couples Dean forces with inertial lift to order the cells prior to encapsulation, as illustrated by images in (B). The periodicity of the regularly-spaced cells can be matched to that of the drop formation, to fill most drops with single cells (C). Reproduced from [77] with permission of The Royal Chemical Society.

An alternative ordering strategy that circumvents some of these issues is close-packed encapsulation[78]: When regularly-sized particles are packed together, they spontaneously organize into a lattice that minimizes packing forces. The periodicity of the lattice can be used to generate

periodic particle flow into a drop maker, to achieve high encapsulation efficiency. Because this approach does not require inertial effects, it can be used over a flexible range of flow rates, making it easier to integrate into devices incorporating multiple microfluidic components. However, it is yet to be demonstrated with cells, which may be difficult to pack and prone to aggregation.

Another strategy for achieving high encapsulation efficiency is cell-triggered jet breakup[79]. In this approach, cells are introduced into a drop maker randomly and used to trigger the breakup of a jet. When no cell is present, very small drops bud from the end of the jet; however, when a cell flows into the jet it perturbs the jet, triggering the formation of a larger-than-average drop with the cell encapsulated inside. The result is a bi-disperse emulsion consisting of small and empty drops and large drops containing single cells. The cell-containing drops can then be selectively recovered from the emulsion using passive methods like pinched flow fractionation[80]. This encapsulation approach may potentially be the most valuable for biological applications of these techniques because it does not require high flow rates or close packing of the cells. However, it is yet to be widely adopted, possibly due to the complexity of integrating the cell triggering and sorting devices into a microfluidic system, which may not be worth the gain of the increasing the fraction of usable drops

The encapsulation of biomolecules into drops is achieved exclusively using limiting dilution because molecules are too small to passively organize with inertial or pack methods and or to trigger the breakup of a jet. By tuning concentration and drop size, it is possible to precisely set the fraction of drops containing single molecules[81–83].

2.5 Further processing of droplets

After the droplets have been formed and loaded with the cells or biomolecules to be analyzed, several additional operations may be needed to complete the analysis, such as adding

reagents to them, incubating them at controlled temperature, or sorting them to recover the ones most interesting for further study. In this section, we describe the techniques that have been developed for processing drops that are most suitable for ultrahigh-throughput biological studies.

Adding reagents to surfactant-stabilized drops

Next to droplet generation, reagent addition is perhaps the most important operation for ultrahigh-throughput applications. Fundamentally, reagent addition allows the execution of multistep reactions in the drops, allowing the drops to be loaded with specific reagents in one step, incubated, and then the conditions in the drops changed by adding another set of reagents.

The challenge with reagent addition is that the drops are almost always stabilized by surfactants, which are essential for preventing them from coalescing upon contact and thus retaining their integrity as separate microreactors. Reagent addition techniques temporarily destabilize the drops so that the reagents can be added without merging the drops. The techniques that are available can be grouped into two categories, passive and active reagent addition. Whereas passive methods allow reagents to be added without the use of external forces, active methods use external forces like electric fields and focused laser beams.

One passive method uses a "push-pull" chamber that expands and then contracts; when two drops flow into the chamber, they merge as they exit through the contracting region[84]. This study demonstrated that, unexpectedly, droplet coalescence is favored by pulling the drops apart rather than squeezing them together. Alternatively, surfactant stabilized drops can also be fused with drops that are not stabilized by flowing them through a zigzag geometry. This strategy enables efficient one-to-one, two-to-one, and three-to-one droplet fusions[85,86]. These passive techniques have the advantage of being exceptionally simple to implement, requiring no specialized electrodes

or laser beams to be integrated into the device, as is needed in active methods. However, a disadvantage of these methods is that the fusion depends sensitively on the chemical properties of the droplets and the dimensions and flow conditions in the microfluidic device; this can make achieving efficient fusion for all biological reagents challenging, particularly when surface-active compounds, like detergents, are present in the drops.

A passive method that overcomes this issue is wettability patterning, in which a hydrophilic strip is patterned into a microfluidic channel[87]. When pairs of droplets flow over the strip, they wet the wall and coalesce with each other and eventually re-form into a single droplet that travels downstream. This method is more robust with respect to the kinds of droplets that can be fused, but has a tendency to transfer material between successive pairs of drops, since material left behind by one pair can be absorbed by a later pair.

Active methods are more difficult to implement because they require the integration of specialized components into the device, like electrodes, but are also more robust in operation and can be applied to a wider variety of reagents[88]. Active methods employ external forces, such as an electrical field, to induce droplet merger. Heat and light have been used for droplet merger[89] but electrical methods have the best potential for ultrahigh-throughput applications because they can merge droplets very quickly and, thus, applied to drops moving at high velocities past the electrode region. One approach uses electrodes submerged in the reagents to create pairs of drops with opposite charge, forcing them to coalesce[90]. A more robust method uses electric fields not in contact with the reagents to create an electric field that the drops must flow through as they move through the microfluidic device. The electric field causes the conductive drops to temporarily polarize, leading to drop-drop interactions that induce coalescence[88,91,92], as shown in Figure 2.4A.

In both passive and active droplet merger techniques, it is imperative to synchronize the trains of drops to be merged so that one of each type of drop is paired together and the pair merged. This is often achieved by making one of the drops smaller than the other. In Poiseuille flow, the parabolic flow profile in the channel causes the small drops to flow faster than the large drops so that they tend to catch up to the larger drops, forming pairs, at which point they can be fused[91,93–95]. For pairs of similarly sized droplets synchronization can be achieved by incorporating expansion chambers to slow down the leading drop[84,96] or electric fields[97,98] to trap the leading drop, allowing the lagging drop to catch up so that the pairs fuse.

A method that bypasses the need for synchronization is picoinjection. In this approach, the droplets to be injected are flowed past a channel containing the reagent to be added. As they pass the channel, an electric field is applied rupturing the surfactant-stabilized interfaces and allowing the reagent to enter the drop, as shown in Figure 2.4B. This approach allows injection of surfactant-stabilized drops and can perform single and combinatorial injections at kilohertz rates[99].

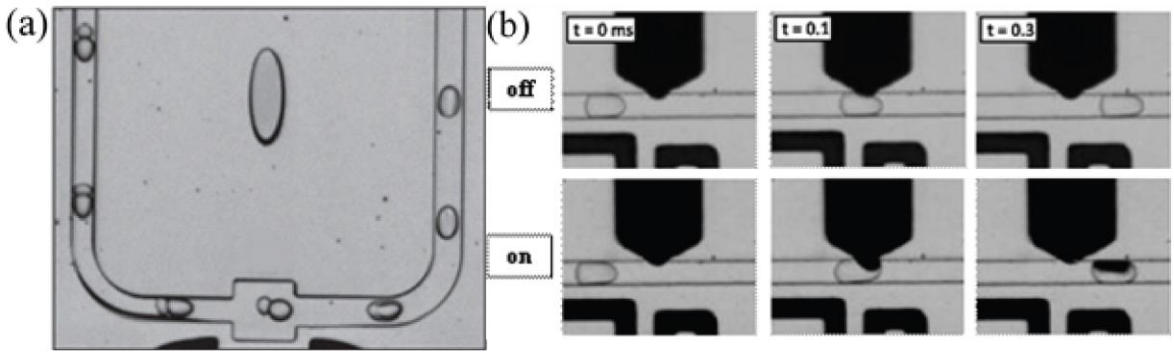


Figure 2.4 The addition of reagents to already-formed droplets is essential when performing most biological assays, since most assays require the addition of different reagents to the drops at different times. The most common techniques for doing this are to coalesce the target drop with a drop of the reagent to be added, (A), or to inject the reagent using picoinjection, (B). Both strategies

exploit electric fields to temporarily destabilize the droplets so that the target drops merge with the reagent to be added. (A) Adapted with permission from Macmillan Publishers Ltd: Nature Biotechnology 91, copyright 2009, and (B) reproduced with permission from 99 .

Rapidly mixing the contents of drops

Due to the small dimensions and relatively low flow rates utilized in most microfluidic studies, the conditions in the channels are laminar; turbulence is absent and streamlines follow paths dictated by the channel geometry. A consequence of laminar flow is that mixing is slow compared to the convective motion of the fluids, which often necessitates the addition of other strategies to enhance mixing. Droplet-based microfluidics is an exception in which mixing is not normally a concern. Even though droplet-based systems operate under laminar flow conditions, mixing is rapid due to the natural generation of recirculating flows in the drops, a result of the drag of the channel walls on the edges of the drops. The recirculating flows decrease the striation length – the average distance for which mixing occurs via diffusion between two materials – so that mixing via diffusion happens much more rapidly than in single-phase fluidic systems[100,101]. Mixing can be enhanced further by flowing droplets through a zigzag geometry, which causes the recirculating flows to change directions, increasing the exchange of fluids between the two halves of the drop and greatly accelerating mixing (Figure 2.5)[102,103]. This geometry can mix drops in milliseconds, fast enough for most ultrahigh-throughput applications[50,104].

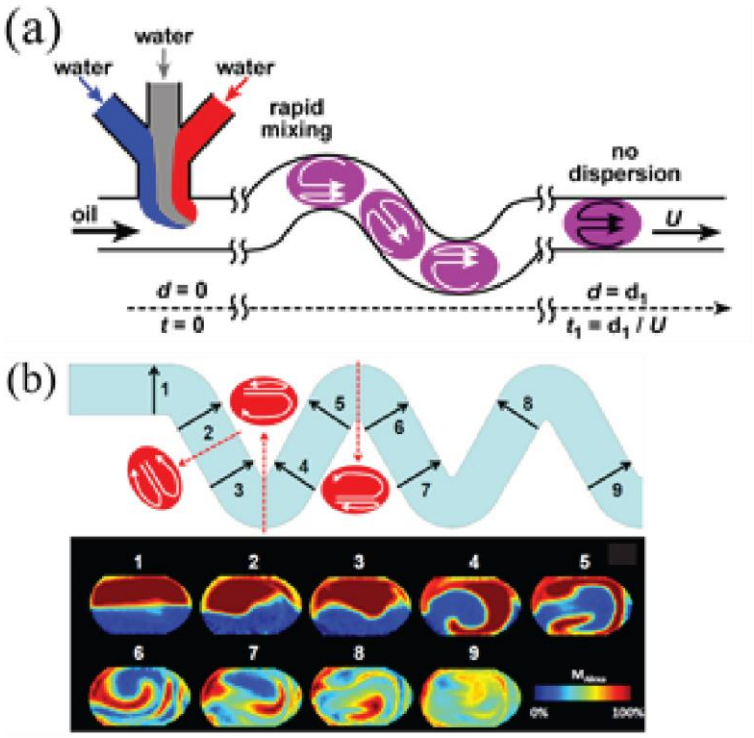


Figure 2.5 Mixing in droplets can be accelerated significantly by flowing the droplets through curving channels. (A) Illustration of streamlines within droplets passing through a curving channel. (Reproduced with permission from [104]). (B) Images of two dyes mixing in a serpentine channel with droplet positions in channel indicated by numbers (Reprinted with permission from [103]. Copyright 2012, American Institute of Physics).

Recovering the contents of drops

In some implementations of droplet-based microfluidics, the final operation is to optically scan the drops to characterize the outcomes of the reactions. However, there are other instances in which the material in the drops must be recovered for further analysis. The simplest way to achieve this is to coalesce all the drops to form a single aqueous pool. Several methods have been developed to coalesce drops, including flash freezing[105] or by adding a chemical destabilizer to the emulsion[82,91,106].

In other instances, in addition to coalescing the drops, it is desirable to ensure that the drop contents remain distinct after coalescence. This can be accomplished by solidifying the interior of the drops prior to coalescence using a gelling agent, such as alginate or agarose. After the drops coalesce, the gelled interiors remain distinct and can be collected and suspended into an aqueous carrier solution. This has been used, for instance, to perform PCR analysis on single cells[107,108]. Functionalized beads can also be used to selectively capture specific analytes from the drops, such as a target protein or DNA molecule[82,109–112]. The beads can be recovered by breaking the emulsion and using a separation technique, like magnetic separation or centrifugation[111].

The droplet contents can also be recovered by selectively fusing them with an aqueous stream using a microfluidic device, Figure 2.6. The droplets pass an opening containing a flowing stream of aqueous fluid. Due to surfactants in the oil phase, in the absence of other forces the droplets do not merge with the stream and exit the device intact. However, when an electric field is applied the droplets fuse with the stream, releasing their contents into it for collection[95,113]. This process is very high throughput and has demonstrated the capability of selectively fusing drops at rates above ten kilohertz.

Another strategy for recovering material from drops that does not require the extraction of the entire drop is geometrically-mediated splitting[114]. In this approach, the drops are flowed into a channel that bifurcates into two channels; as the drops flow through the bifurcation, they are divided into two portions, one traveling down each arm of the split. The volumes of the two portions and, thus, of fluid that is sampled from each drop, can be adjusted by controlling the flow rates through the two arms, either by pressurizing the outlets or tuning hydrodynamic resistances[67]. The different portions of the droplets can then be analyzed by optically scanning them or merging them with droplets containing other reagents[115–117].

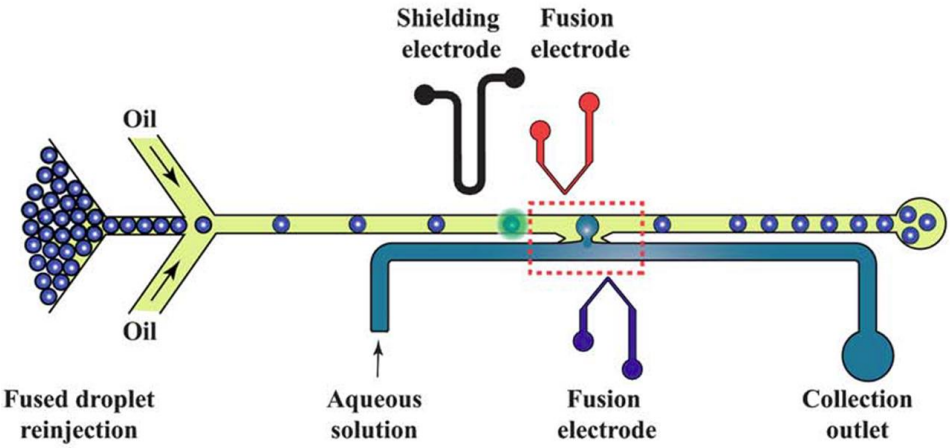


Figure 2.6 Electrocoalescence of drops with an aqueous stream for recovering target drops. Reproduced in part from [95] with permission of The Royal Society of Chemistry.

Incubation and storage of drops

In nearly all biological assays, steps of thermally-controlled incubation are required, such as to stimulate the growth of bacteria or perform PCR amplification of nucleic acids. In droplet-based microfluidics, droplets can be incubated off the microfluidic device, for example, in an incubator or PCR machine, or on the microfluidic device, using integrated thermoelectrics and temperature controllers[118,119].

On-chip incubation is normally achieved by flowing the drops through channels or trapping them in static arrays. With channel incubation, there are different options depending on the duration of incubation required. For example, for very short incubations (seconds), the drops can be flowed single-file through a channel of controlled length[120]. This has the benefit of keeping the drops in order, but is limited in the duration of incubation it can achieve because long incubations require long channels, which have high hydrodynamic resistances and require impossibly large input pressures to drive the flow. This can be avoided by stopping the flow so that the drops are

static in the channel but, due to the small volume of even a relatively long microchannel, this strategy is only applicable to storing thousands of drops[121].

When longer delays are required, the most common strategy is to use a delay line, which is a wide and tall channel that is able to store a large number of drops. Due to the large diameter of the delay line, the pressure required to pump the drops through it is small, making it appropriate for hours of incubation. The drawback to this approach, however, is that the drops pack into the delay line in three dimensions and, consequently, lose their original order. The drops can also move with respect to one another, resulting in dispersion of the incubation time – that is, in which the duration that each drop spends in the delay line varies. Time dispersion can be overcome using several strategies, including by packing the drops densely so that the drops jam and are prevented from changing positions, as shown in Figure 2.7A[122]. Another strategy is to include periodic constrictions into the line, which cause the drops to repeatedly shuffle; this averages out variations in droplet velocity, causing the incubation time of each drop to converge to the average value, as shown in Figure 2.7B[122].

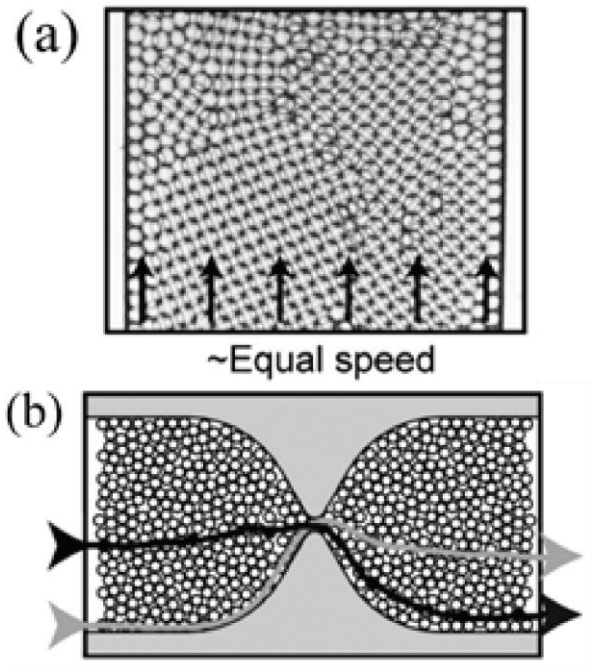


Figure 2.7 Strategies for preventing dispersion of incubation time in delay lines include (A) densely packing drops or (B) implementing repeated constrictions to average out positional-dependent differences in droplet velocity, yielding equal incubation time for all drops. Reproduced in part from [123] with permission of The Royal Society of Chemistry.

Reservoirs are very large channels (millimeters in width) that can store millions of droplets for hours of incubation. Once droplets are packed into the reservoir, the flow can be stopped and the device can be incubated under the desired conditions[24,33,51]. The drops can be packed as a monolayer to allow each droplet to be directly visualized[124]. In addition, if stored in a gas-permeable PDMS device, gas exchange through the channel walls can be used to enhance the survival or cells compared to storage in an air-tight syringe[75].

An alternative strategy for storing drops in a chamber is to use microfabricated features to position the drops at defined locations; this can aid visualization and prevent the drops from moving over the course of the experiment[39,45,125], as shown in Figure 2.8. A limitation of static

arrays, however, is that they are difficult to fabricate with the capacity needed to make them valuable for ultrahigh-throughput applications: While the largest demonstrated array held 1 million drops[126], most have only been able to hold tens of thousands[39,45,118,125,127–131].

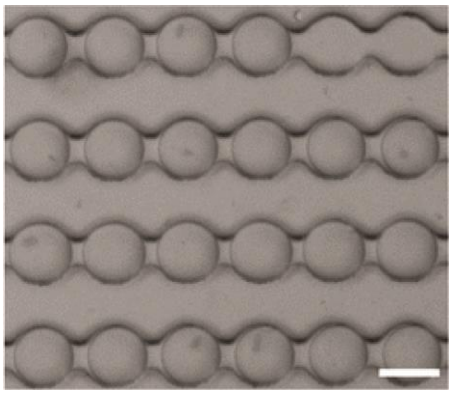


Figure 2.8 Drops stored in a static array of microfabricated chambers. Adapted from [39] with permission of The Royal Society of Chemistry.

When there is not an explicit need to incubate droplets on the microfluidic device, off-chip incubation is often the easiest strategy. Off-chip incubation is often accomplished by transferring the droplets from the microfluidic device into a syringe or centrifugal tube, and then storing the tube at the desired temperature and atmospheric conditions in a cell incubator, thermal cycler, or other controlled environment[34,35,42,132–134]. Off-chip storage is appropriate for hours to weeks of incubation and the emulsion can be re-injected into a microfluidic device for analysis and further manipulation[34]. Long-term off-chip storage of droplets for screening applications has also been proposed using microfluidic cartridges[135,136]. Ultimately, the limit to the duration over which the emulsion can be stored depends on the stability of the emulsion and viability of the reagents or living organisms contained within the drops.

Sorting

The sorting of an emulsion, the selective retrieval from an emulsion of a population of drops, is critical for a number of especially important ultrahigh-throughput studies. It is needed, for instance, when performing directed evolution, in which a large library of drops is prepared, each containing a distinct mutant of a cell or protein, and a small fraction of which are desirable for further study. To perform the experiment, all droplets must be scanned and the ones with desirable mutants must be recovered. Because ultrahigh-throughput biological systems typically operate at droplet rates in the kilohertz, exceptionally fast sorting methods are required.

Ultrahigh-throughput sorting of droplets can be achieved using passive and active means[95,137], although the active method is generally the most useful because it is versatile and can be used to sort droplets based on a complex logical decision. Active droplet sorting has been demonstrated using magnetic[138], mechanical[139], acoustic[140], electrophoretic[90] and dielectrophoretic[141–143] forces to separate the drops. The general approach involves flowing the drops single file into a junction with two or more outlet channels. The junction is designed so that, by default, all droplets flow into one of the outlet channels, termed the "waste," due to differences of the channels in hydrodynamic resistance. Select drops can then be sorted into the other channel by applying one of the abovementioned forces. Of the techniques developed, pressure-based, acoustic, and dielectrophoretic sorting have proven to be the fastest, capable of sorting droplets at rates faster than 200 Hz. Dielectrophoretic droplet sorting has been used to screen a large library of enzymes at rates greater than 1 kHz[142,143], as illustrated in Figure 2.9. An alternative approach to microfluidic sorters is to perform the reactions in water/oil/water double emulsion droplets which can then be scanned and sorted using flow cytometry[144]; however, the lower stability of the double emulsions, coupled with a larger solubility of encapsulated molecules in the aqueous carrier phase, can limit the utility of this approach[142].

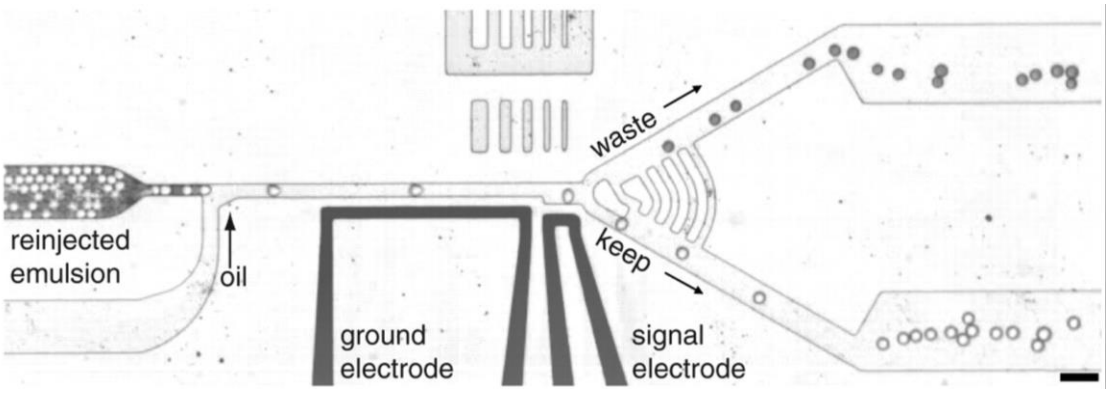


Figure 2.9 Dielectrophoretic sorting of drops at kilohertz rates. Reproduced with permission from [143].

Device integration

The microfluidic components described above can be combined into single, fully-integrated systems designed to perform complex assays[123,145]. However, when integrating multiple components together, it is imperative to ensure that all components can be operated under the same flow and pressure conditions. If a drop maker is to be combined with a sorter, for example, then the drop maker must form drops of the same size and rate needed for the sorter to function. Such optimization can be tedious and require significant investments of time[146]. One way to avoid these difficulties is to use each component individually as its own device[91,95,143]. For example, to form and then sort drops, the fluids for drop formation can be injected into a drop making device at the flow rates needed to form drops of the desired size. The drops can then be collected into a syringe and, after they have all been formed, injected into the sorting device at the flow rates needed for this device. In this way, many operations can be performed in sequence, without having to engineer the components to operate simultaneously, at the same flow rates, on a single chip.

2.6 Labeling and detecting droplets

Labeling strategies

Labeling refers to marking the droplets in an emulsion so that the identity of each droplet can be determined by reading the label at a later time, for instance, to know which reaction is being performed in a given droplet. Methods to label droplets utilize spatial and chemical encoding. In spatial encoding, the identity of the droplet is linked to its location[147], for instance, by positioning drops on an array such that the position of each drop on the array relates its identity[148]. Single-file positioning is limited in the number of drops that can be stored by the length of the channel. Two dimensional arrays allow much larger numbers of drops to be stored and are generally more useful for ultrahigh-throughput studies[149].

Another way to label droplets is to add fluorescent dyes to them. Dyes are can be used to determine which drops have successfully been encapsulated with cells[76,127], cellular components[31], and beads[150]. Dyes can also be used to indicate the concentration of compounds[40,151,152]. Multiple dyes can be used to create "virtual arrays" in which the x-y position of a given droplet on the array is determined by measuring the concentrations of two differently-colored dyes in the drop[40]. Other labeling strategies that are not as widely used include the use of DNA sequences[153,154], quantum dots[155–159], and suspension array technologies[160].

Interrogating drops

Droplets can be interrogated using serial scanning or parallel scanning. Serial scanning interrogates droplets individually as they flow through a specific location in the microfluidic channel. For ultrahigh-throughput applications, this necessitates very high speed interrogation to keep pace with the kilohertz droplet rates. Parallel detection, on the other hand, can utilize slower

measurement strategies, such as CCD imaging, because a large number of droplets can be interrogated simultaneously, yielding a net high throughput. A variety of optical, chemical, and electrical detection methods have been reported for microfluidic biological assays, but few are appropriate for ultrahigh-throughput applications because of their inability to precisely interrogate the drops on the timescales needed to keep pace with the microfluidic devices. Examples of methods with high potential value that have, as yet, not been shown capable of ultrahigh-throughput interrogation of drops are mass spectrometry[135,161,162], electrochemical detection[163,164], Raman spectroscopy[165], surface enhanced Raman scattering (SERS)[166–169], surface enhanced resonance Raman scattering (SERRS)[170], and fluorescence polarization[171].

The most widely used interrogation methods for ultrahigh-throughput applications utilize fluorescence and brightfield microscopy[119,131,145]. Fluorescence enables measurements with high signal-to-noise and spatial resolution, and it can be coupled to a wide variety of existing biochemical assays for which fluorescent readouts already exist. The fluorescence readings can be obtained using high speed, high sensitivity single-point detectors, such as photomultiplier tubes (PMTs)[99,123,127,132,134] or avalanche photodoides (APDs)[76,150,172–177]. The sensitivity and versatility of fluorescence has made this approach useful for a variety of ultrahigh-throughput applications, including cell and protein screening[142,177,178], DNA and protein detection[76,132,150,172], quantitative PCR[105] and dose-response screening[123]. A representative schematic of ultrahigh-throughput fluorescence drop interrogation is shown in Figure 2.10. Fluorescence has also been used to image arrays of droplets for real-time monitoring of large numbers of drops[125]. Fluorescence resonance energy transfer (FRET), another technique that affords information about molecular-scale conformations, has been used to measure binding kinetics[150,176,179], enzyme kinetics[173], and protein-protein interactions[180] at kilohertz rates. Fluorescence lifetime imaging (FLIM) has also

been demonstrated with microsecond temporal resolution for reconstructing mixing patterns within droplets[174,181]. The main challenge with fluorescence-based detection methods is their dependence on a suitable fluorescent marker that can be attached to the molecule of interest. Such tags are not readily available for all molecules and the properties of small molecules can be significantly altered by the presence of an extrinsic marker. Indeed, the primary challenge when performing reactions in microdroplets is identifying a suitable fluorescent marker.

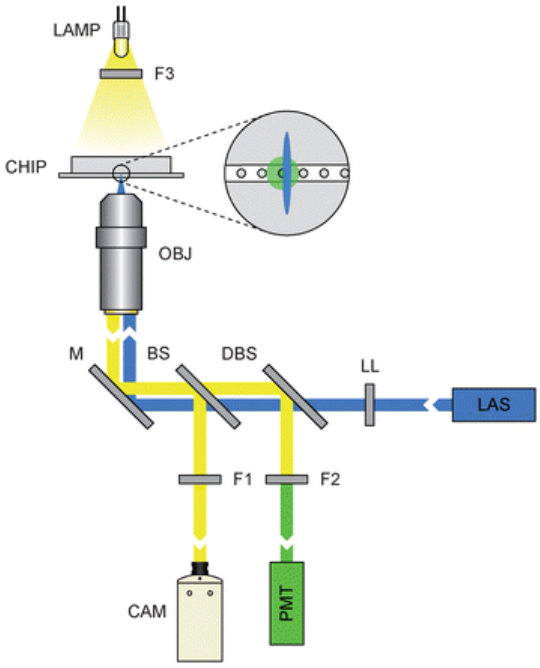


Figure 2.10 Schematic of fluorescence activated droplet detection. Laser light (LAS) is used to excite droplets. The fluorescent emission from each drop is detected with a photomuliplier tube (PMT). Filtered light from a halogen lamp (CAM) allows droplets to be monitored with a high speed camera (CAM). Reproduced from [142] with permission of The Royal Society of Chemistry.

2.7 From Tubes to Drops: applications of droplet-based microfluidics in biology

There are numerous examples in biological research in which the major bottleneck is screening massive libraries of reactions. For example, directed evolution of enzymes, which generate enzymes for a host of biotechnology and industrial applications, often require the screening of libraries of enzymes numbering in the billions of mutants[182]. Small molecule screens for drug discovery often require the analysis of millions of compounds, each of which must be tested in a separate reaction chamber[167]. Such screens are currently conducted using conventional high-throughput methods with pipetting robots in 384 or 1536 well plate formats. With a maximal screening rate of ~1 Hz and a minimal sample volume of 100 nL, screening a million compounds would require weeks of continuous runtime and litres of total reagent, making it impractical for all but the best-funded operations and, even then, for only the most valuable targets[167].

Droplet-based microfluidics affords an exciting and potentially vastly superior methodology for making large library screening accessible by significantly reducing the volume and increasing the throughput of the screening. With this approach, it is possible to screen reactions at a rate of ~1 kilohertz using ~1 pL per reaction, allowing millions of reactions to be screened in minutes using microlitres of total reagents[183]. The barrier to implementing this approach, and realizing its enormous potential for future screening needs, is integrating all of the necessary capabilities into onto one microfluidic workflow. In this section, we describe the first forays into this arena and the use of these techniques to perform novel screens with unprecedented reductions in screening time and cost.

Biological operations in droplet-based microfluidic screening

Most biological protocols involve the repeated transfer of reagents into and out of a reaction vessel, combined with incubation and separation of reactants and products. In theory,

these operations can all be performed in microdroplets using microfluidic devices to perform the various needed operations of reagent transfer, incubation, and sorting.

In addition to these common microfluidic operations, many screening applications also require executing a recurring set of biological operations. One such operation often needed in cell-based studies is to culture cells in drops. Mammalian, algae, and fungal cells have all been cultured in drops for several days[41,42,184]. Another important operation is to transfect cells to induce the expression of a protein of interest or to knock out a gene of interest, which has been achieved in drops using chemical agents[35], electroporation[185], and viral infection[42] at efficiencies comparable to bench top methods. DNA can be extracted by lysing cells in drops[108], and purified from lysate by repeated washing of DNA binding magnetic beads co-encapsulated in the drops[114]. Another valuable operation is to perform entirely in vitro, cell-free expression and characterization of enzymes in drops[24]. This is valuable because existing ultrahigh-throughput screening methods utilizing FACS require a cell that can both synthesize the enzyme and be used to characterize its activity. Host cells may not survive screens that involve the use of toxic substrates or products, or require extreme conditions of temperature or pH.

Early work into the kinetics of enzymes demonstrated that microdroplet encapsulation affords the sensitivity needed to measure the activity of single molecules of β-galactosidase[186]. Similarly, digital emulsion PCR, in which individual molecules of DNA are encapsulated in drops and amplified, has now become a standard biological protocol used in many applications, including as an alternative to quantitative PCR and to prepare DNA libraries for next generation sequencing[112,187]. Other amplification strategies have also been demonstrated with DNA and RNA in drops, including RT-PCR[188] and rolling circle amplification for diagnostic and directed evolution applications[133].

One of the challenges of performing complex biological operations in microdroplets is if the protocol requires a large number of steps, it can be challenging to build a fully integrated microfluidic platform to do this. In these instances, a possible solution is to solidify the drops using a gelling agent, such as agarose[108], so that the emulsion can be broken and the microgels re-dispersed into an aqueous carrier phase. This allows the numerous steps of the protocol to be performed off microfluidic device using standard bench-top methods, while partially maintaining the compartmentalization of the drops. This approach, which has enormous potential for single-cell analysis, has already been demonstrated for performing PCR analysis on large numbers of single cells[23]. However, molecules smaller than the pore size of the gels can exit the gels via diffusion, making this approach only applicable to molecules larger than the pore size or that can be chemically bonded to the gel matrix.

Ultrahigh-throughput studies utilizing co-flowing stream encapsulation

Co-flowing stream encapsulation allows multiple reagents to be combined immediately before being encapsulated in microdroplets. Due to laminar flow conditions in the channels, the reagents in the different streams do not mix until after they are in the drops. This is useful for a variety of applications of ultrahigh-throughput studies. For example, this approach has been used to profile a population of bacteria for sensitivity to different antibiotics. The different antibiotics were first loaded into the channel as long plugs separated by plugs of an inert spacer fluid. The plugs were then merged with reagents needed for the assay and the bacteria and immediately encapsulated in drops using a T-junction drop maker, as shown in Figure 2.11. This approach allowed multiple antibiotics at several concentrations to be rapidly screened[189]. The same principle was applied to screen the effects of varying salt, protein concentration, and precipitants to identify

optimal parameters for protein crystallization[145,190]. This principle was further applied to generate high-resolution drug dose-response curves[123].

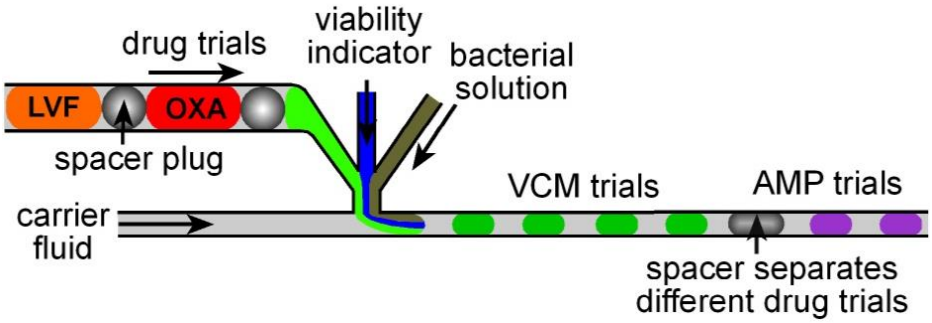


Figure 2.11 Co-flow drop formation used to assay drug susceptibility of bacteria. Figure adapted from [189].

Ultrahigh-throughput screens using combinatorial droplet merger

Co-flowing stream encapsulation is suitable for screening a small number of reagents at different concentrations and, combined with a reagent-plug technique like the one described above, can also be used to screen moderately large sets of compounds, numbering in the tens-to-hundreds. However, for larger numbers of compounds, it becomes impractical because it requires very long tubes of plugs or large numbers of inlets on the microfluidic device. In these cases, a superior approach is combinatorial droplet merger. The basic strategy of combinatorial droplet merger is to use an automated method to create a "droplet library" of the different compounds to be screened, as shown in Figure 2.12A; the droplet library consists of an emulsion of monodisperse droplets in which each droplet contains a different reagent or concentration of reagent. The droplets library can then be merged with other droplets using electrocoalescence, as illustrated in Figure 2.12B. This can be useful for protein crystallography, where screening for optimal crystallization conditions is a bottleneck[191,192].

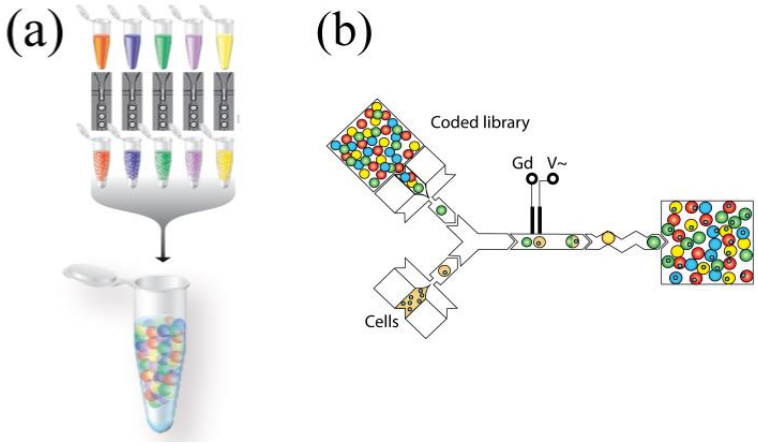


Figure 2.12 Creation and use of a drop library for combinatorial screening. (A) Multiple reagents are emulsified and pooled, forming a droplet library. (B) The library of reagents are randomly paired and merged with the assay drops. Adapted with permission from [40].

A key consideration when using combinatorial droplet merger to screen a library of compounds is indexing each drop to keep track of which compound is being tested. One way to do this is to label the droplets with fluorescent dyes, as described in the previous section of droplet labeling, in which fluorescent dyes of different color are loaded into the droplets at different concentrations. Each colored dye can be used to represent a digit in a multi-digit number, where the concentration of the dye sets the value of the digit. An approach like this was used to screen a drug library of mitomycin C at different concentrations to characterize cytotoxicity[40]. However, optical labeling of drops with fluorescent dyes is limited by the finite spectrum of visible light and the precision with which dye concentrations can be measured in the drops.

Another important application of combinatorial droplet merger is to perform highly multiplexed targeted PCR amplification of genomes. The genomes of organisms are often massive, containing millions or billions of base pairs of information, but the information that is the most relevant for study may be localized in specific regions. The challenge is to target these regions for

sequencing and discard the vast "background" of genomic DNA in the other regions. One way to do this is to use multiplexed PCR, in which primer sets are designed that target each of the desired regions for amplification. This allows, in principle, only these regions to be copied, discarding the rest of the genomic DNA. However, in practice it can be extremely challenging to optimize multiplexed amplification of so many regions, limiting the total number of regions that can be simultaneously interrogated. In these cases, emulsion PCR with combinatorial droplet merger affords an excellent solution[91] where each of the PCR amplifications for the different primer sets are performed in a separate microdroplet. Because the primer sets are isolated from one another, they do not interact or compete for binding to the genomic DNA. This yields much cleaner PCR products and allows more uniform amplification of all regions. A technology based on this approach is in fact already being commercialized by RainDance Technologies as a pre-sequencing front-end to enrich for genomic regions of interest.

Ultrahigh-throughput applications utilizing droplet sorting

Droplet sorting is essential in applications where specific droplets must be recovered from a large population of other drops. To date, the majority of microfluidic studies utilizing droplet sorting have been focused on performing protein engineering through directed evolution, a technique in which proteins are evolved to enhance a desired property by processing them through successive rounds of mutagenesis and screening. Droplet-based methods of directed evolution are used because the drops can encapsulate products that are secreted from cells as well as provide encapsulation for cell-free systems. This is particularly important for evolving enzymes, which typically catalyze reactions in the liquid phase and generally release products into solution. Methods utilizing FACS cannot easily accomplish this because FACS requires that that the product

of the enzyme be localized within the cell. While certain enzymes can be screened inside cells, for many enzymes this is not possible.

Droplet-based microfluidic screening thus holds enormous potential as a general platform for evolving enzymes that, presently, can only be evolved using low throughput methods, like screening on a well plate array. At a screening rate of 2 kHz, more than 108 droplets can be sorted per day, matching the library size that can be expressed using yeast. Compared to 384 well plates, droplet screening achieves 3-4 orders of magnitude greater throughput and ~6 orders of magnitude reductions in reagent costs[143].

The principle strategy in directed evolution with microdroplets is to isolate each variant of the protein in a different droplet, screen all variants, and recover the best by sorting. The precise mechanism by which the protein is synthesized from the gene varies and can be accomplished using bacteria, yeast, or even cell-free expression methods. Directed evolution with droplets has already been used to improve the catalytic rate of horseradish peroxidase[143], to screen a retrovirus display library for active tissue plasminogen activator[178], and with an in vitro transcription-translation system to select for highly active β-galactosidase from a mixed population[95].

Beyond directed evolution, droplet sorting has also been used as a post-enrichment screen for aptamer libraries, to replace the costly sequencing and re-synthesizing steps of the SELEX workflow for aptamer discovery[193]. The secretion of IL-10 by immune-suppressing cell populations in the bloodstream was detected using antibody functionalized beads co-encapsulated with the cells in drops. Sorting based on IL-10 secretion can be used to enrich for rare cells previously isolated[194].

Digital assays on single molecules or cells

For a fixed number of molecules, reducing the volume of the reaction vessel effectively increases the concentration of the molecules. By shrinking the reactor to the size of a microdroplet, the effective concentration of a single molecule per drop can be comparable to the concentrations used in bench top assays. This concept was employed in the early 1960s to measure the activity of individual molecules of β-galactosidase encapsulated in drops, as shown in Figure 2.13A[186].

Microdroplet encapsulation is thus one way to significantly increase the sensitivity of an assay. Another major advantage is that it can isolate molecules in a complex mixture so that they can be interrogated individually. For example, single molecule encapsulation has been combined with multiplexed PCR to identify and quantify multiple DNA targets in an originally mixed population[195]. Microdroplet "digital" quantification of encapsulated molecules affords an alternative strategy to quantitative PCR that provides absolute molecule counts without the need for normalization or calibration (Figure 2.13B)[196]. Digital PCR can achieve higher sensitivity and dynamic range than standard quantitative PCR methods[197,198]. Digital PCR has been used to detect aneuploidy with precision that exceeds what is possible with other methods and has been successfully implemented into a commercial laboratory instrument for DNA quantification[133,198,199]. Using a similar approach to digital detection of nucleic acids, rare cells can also be detected "digitally" with microdroplet encapsulation[23]. For further information on digital PCR in droplets, we refer the interested reader to this recent comprehensive review[200].

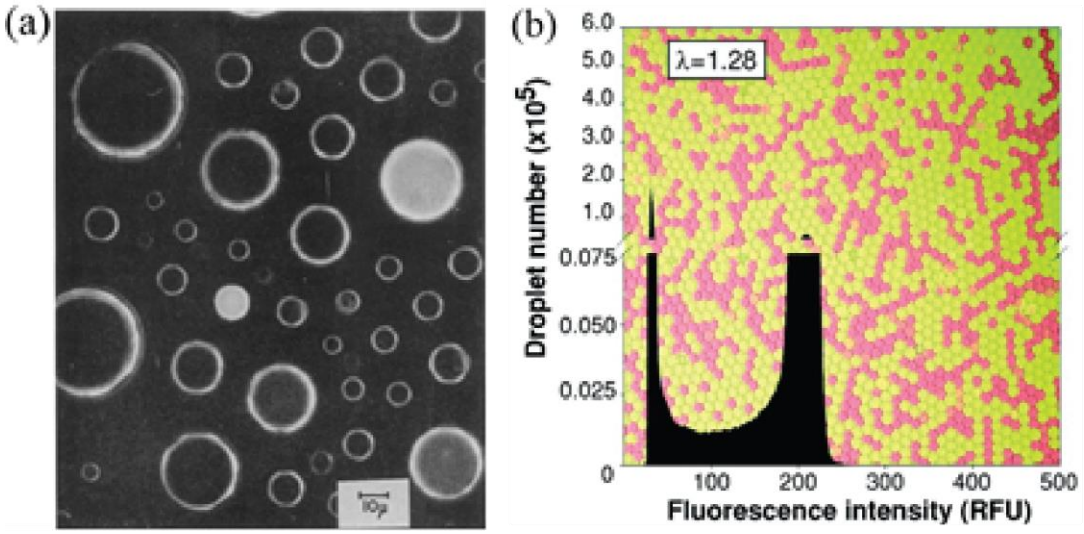


Figure 2.13 Single molecule analysis in drops. (A) The activity of single molecules of β-galactosidase can be measured using microdroplets, as illustrated by the fluorescently bright drops (Reproduced with permission from [186]. Copyright 2005 National Academy of Sciences, U.S.A.). (B) Amplification of single DNA molecules in microdroplets using a DNA intercalating green fluorescent dye and dextran-Texas red as a reference. Adapted with permission from [196]. Copyright 2009 American Chemical Society.

2.8 Conclusions

The field of droplet-based microfluidics is exploding and has yielded numerous technologies that are enabling for applications throughout biology. The next critical steps in the development of this field are for microfluidic research labs to continue to develop robust and widely applicable tools and to demonstrate the power of these tools through proof-of-principle experiments that biologists will recognize. In addition, collaborations with industry will be crucial, because like all high technologies droplet-based microfluidics will only be adopted widely when systems are built that are easy to use, cost effective, and robust in operation, even with operators that are not familiar with the underlying microfluidic technology. This is a tall order but it has been

achieved before, as evidenced by the numerous high technologies commonly use in biological research labs, like next generation sequencing platforms, mass spectrometers, and FACS machines. Indeed, FACS, in many ways, is the predecessor of droplet-based microfluidics and affords an excellent example of what we, as a field, should strive for in the translation of the technology: Like droplet-based systems, FACS utilizes microfluidic channels, high-speed optical detectors, and high speed computational analysis for data capture and logic-based sorting. We envision a time in the near future when FACS machines will be upgraded with droplet-based systems to yield a much more general screening platform – one that is not limited to the screening of cells alone, but can conduct, with full automation, and screen any variety of liquid-phase reactions, involving cells, biomolecules, and other reagents. In addition, while in this Review we have focused on the biological applications of this technology, the same capabilities of high efficiency, ultrahigh throughput, and precision of measurement will make it valuable for screens in other fields too, particularly in the chemical sciences.

Chapter 3: Adhesive-based bonding technique for PDMS microfluidic devices

The following section is reprinted from "Adhesive-based bonding technique for PDMS microfluidic devices" by C. Shea Thompson and Adam R. Abate. The manuscript was published as a Communication in Volume 13 of Lab on a Chip. C. Shea Thompson and Adam Abate designed the experiments and wrote the manuscript. C. Shea Thompson performed the experiments.

In this chapter, we describe a method for bonding PDMS microfluidic devices that uses only adhesive tape and an oven. The simplicity, low cost, and reproducibility of our method should allow it to be adopted by researchers lacking access to specialized equipment or cleanroom facilities. This technique also allows devices to be bonded and used within two hours, which is convenient for rapid prototyping and testing. As interest in microfluidic methods moves beyond specialized engineering laboratories, there is a need for alternative PDMS bonding methods that can be readily adopted by researchers who are not specialists in microfluidics.

3.1 Abstract

We present a simple and inexpensive technique for bonding PDMS microfluidic devices. The technique uses only adhesive tape and an oven; plasma bonders and cleanroom facilities are not required. It also produces channels that are immediately hydrophobic, allowing formation of aqueous-in-oil emulsions.

3.2 Introduction

Soft lithography in polydimethylsiloxane (PDMS) is a ubiquitous method for rapid prototyping of microfluidic devices[201]. A critical step in the fabrication involves sealing the devices by bonding the PDMS channels to a substrate. Numerous PDMS bonding strategies have been reported[28,202–204], but oxygen plasma treatment is the most common. While oxygen plasma is

effective and produces strong bonds, the necessary equipment is expensive and access to cleanroom facilities is limited. Other methods, like partial cure bonding or the use of chemical crosslinkers, can also be used to bond devices, but often require hours, if not days, to complete the bond before the devices can be used. As interest in microfluidic methods moves beyond specialized engineering laboratories, alternative techniques for bonding PDMS devices will be useful. In particular, methods that are simple and inexpensive will enable the broadest adoption of these techniques by researchers in other fields.

We present a simple and inexpensive bonding technique for PDMS devices that requires only adhesive tape and an oven. Adhesive tape is applied to the bottom surface of the PDMS device and the device baked at 65°C for 2 hours. The baking increases the bond strength to the tape, allowing the devices to support pressures of tens of kilopascals for hours of operation. A variety of common adhesive tapes can be used, including optically transparent tapes that enable brightfield microscopy of the channels and double-sided tapes that can be adhered to another substrate, like a rigid glass plate, providing even stronger bonds. We demonstrate the biocompatibility of the method by using a tape-bonded device to generate droplets for emulsion PCR. The simplicity, low cost, and reproducibility of our method should allow it to be adopted by researchers lacking access to cleanroom facilities.

To test the burst strength of PDMS devices bonded with tape, we fabricate smooth PDMS slabs with holes punched into them. PDMS elastomer (Sylgard) is prepared by mixing the elastomer base with crosslinker at a 10:1 weight ratio using a Dremel hand drill. The mixture is degassed under vacuum for 30 minutes, poured into a plastic Petri dish, and cured for 2 hours at 65°C. The cured PDMS is sectioned into 3 x 3 cm^2 slabs using a razor blade. A 0.75 mm diameter tubing inlet hole is cored into the center of the device (Harris Unicore). The slab is washed with

isopropyl alcohol and dried with compressed air. One of three adhesive tapes (Scotch®MagicTM Tape, Scotch® Permanent Double Sided Tape, Scotch® MultiTask Tape) is then applied to the slabs and the slabs are baked at 65°C for 0, 1, 2, 4, or 16 hours. Specifications for the adhesive tapes studied are described in Table 3.1.

To produce microfluidic devices with this approach, PDMS replicates are molded from an SU-8 master using the techniques of soft lithography[201]. The device is punched with inlet ports, washed, tape-bonded, and baked, as illustrated in Figure 3.1.

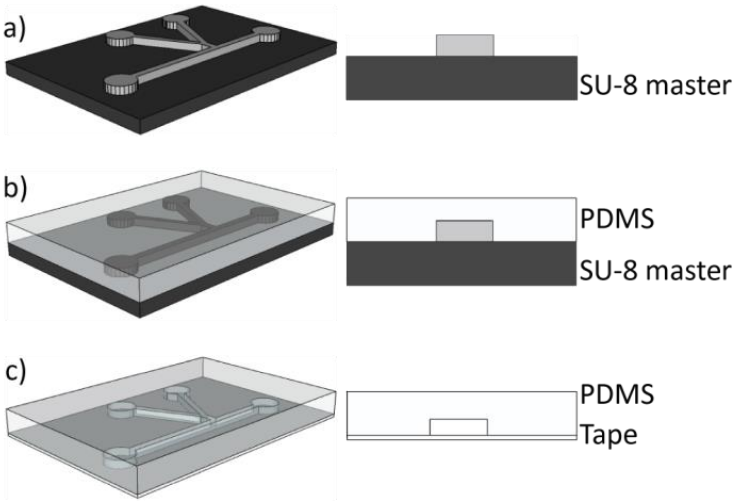


Figure 3.1 Fabrication process for bonding PDMS with adhesive tape. (A) SU-8 master is fabricated on a silicon wafer using photolithography. (B) PDMS is cast on the master and cured. (C) The PDMS replicate is removed, punched with inlet ports, washed, and adhesive tape is applied to its bottom surface. The device is baked at 65°C for 2 hours to complete bonding.

Microfluidic bonding techniques must produce strong bonds that prevent fluid leakage at the PDMS-substrate interface. To measure the strength of the PDMS-tape bonds and identify the optimal bake time for strengthening the bonds, we use burst pressure testing. Polyethylene tubing is inserted into the inlet port of the tape-bonded PDMS slabs, through which regulated air pressure

is applied. We increase the air pressure in increments of 7 kPa at time intervals of 30 s until the PDMS-tape bond breaks.

We monitor the bond integrity under a brightfield microscope to determine when the bond begins to fail. To verify there are no air leaks at the device inlet, we monitor a soap-water solution applied at the device-tubing interface.

3.3 Results

The PDMS-tape bond is stable up to pressures at which Saffman-Taylor fingers form around the inlet[205,206]. Saffman-Taylor fingers are characteristic of the interface that forms when a viscous fluid, such as the adhesive from the tape, separates between two diverging surfaces. As the applied pressure increases, the Saffman-Taylor fingers continue to develop, up until the point that the bond fails, as depicted in Figure 3.2A.

We test the bond strengths of three commonly available adhesive tapes, a comparison of which is provided in Figure 3.2B. For each column, the lower value represents the pressure at which Saffman-Taylor fingers begin to develop and the higher value the pressure at which the bond fails. The bond strength increases with baking time from 0-2 hours for all adhesive tapes but, for baking times greater than 2 hours, does not increase. The double-sided tape reproducibly yields the strongest bonds, and bonding the bottom surface of the double-sided tape to a rigid glass slide can increase the bond strength further. The maximum bond strength achieved after 2 hours of baking is comparable to that of oxygen plasma and corona discharge[207], but the PDMS-tape bond is not permanent and fails after minutes to hours under the maximum applied pressure. These data

show that the bond strengths obtained with common adhesive tapes are sufficient for many microfluidic applications.

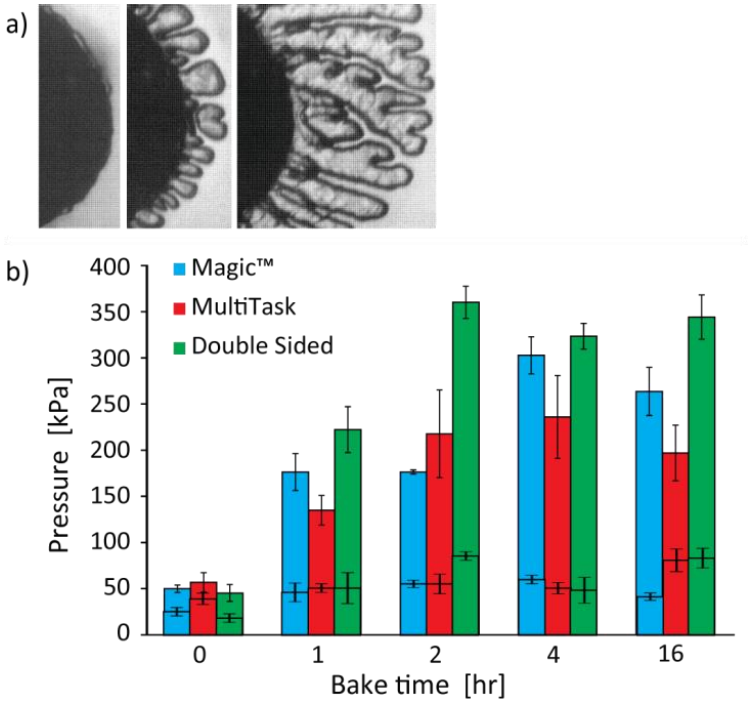


Figure 3.2 PDMS-tape bond strengths. (A) Brightfield microscopy image of the formation of Saffman-Taylor fingers as pressure is applied to a PDMS-tape bond through an inlet port; the inlet port is a circular hole, the right edge of which is visible in the left of the images. The pressure increases from left to right. (B) A comparison of bond strengths for three adhesive tapes baked for different times. For each bar, the lower value is the pressure at which Saffman-Taylor fingers form and the higher value the pressure at which the bond fails. Each bar is the average of three experimental measurements with different devices and error bars denote the standard error of the mean of these values.

An important consideration when fabricating microfluidic devices for droplet-based applications is producing channels with the desired hydrophobic wettability, so as to allow the

formation of aqueous-in-oil emulsions. Commonly used methods like oxygen plasma bonding or chemical bonding often render PDMS hydrophilic[208], necessitating additional steps of processing to regain hydrophobicity. These include baking the devices for long durations to allow the PDMS to revert to its native hydrophobic state[209] or functionalizing the surfaces of the channels with hydrophobic silanes[210,211] and other chemical modifications like Aquapel[212]. In addition to increasing fabrication time and complexity, these steps are prone to failure, yielding channels with improper wettability and preventing the robust formation of emulsions. By contrast, our adhesive tape bonding method reliably produces channels with the needed hydrophobic wettability and also allows the devices to be used immediately without additional processing steps. This is because the PDMS remains in its native hydrophobic state throughout the bonding process and the adhesive of the tape, which comprises the bottom surface of the channels, is hydrophobic as well.

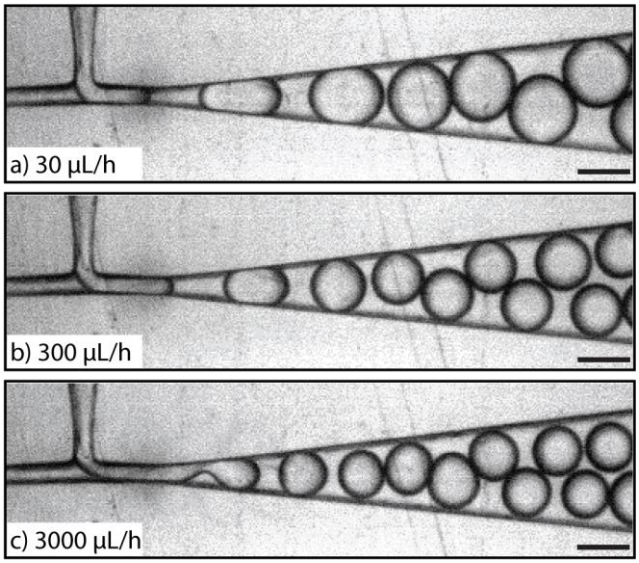


Figure 3.3 PDMS-tape bonded T-junction drop maker used to create monodisperse microdroplets. We vary the flow rates over two orders of magnitude for a fixed oil-to-aqueous fraction of 2:1: Total flow rate (A) 30 µL/h, (B) 300 µL/h, and (C) 3000 µL/h. Scale bars represent 50 µm.

To illustrate that tape-bonded devices have the requisite hydrophobic wettability to form aqueous-in-oil emulsions, we fabricate a T-junction drop maker bonded to Scotch® MultiTask tape. We choose this tape because it is transparent, allowing brightfield monitoring of droplet formation. For the emulsions we inject distilled water into the dispersed phase inlet and the fluorinated oil HFE-7500 with a biocompatible fluorinated surfactant[47] into the continuous phase inlet. These fluids intersect at the T-junction, where drops are formed, as shown in Figure 3.3 for three different flow rates. We vary the flow rates over two orders of magnitude to demonstrate that our adhesive-tape bonding method has the strength needed to operate devices under relevant flow conditions. Indeed, the highest flow rate, depicted at Figure 3.3C, is the maximum at which drop formation in this device is possible; at this flow rate the device no long forms drops in a regular dripping process but rather jets, as can be seen by the fluid tongue extending into the expansion along the upper wall of the T-junction. This device behaves essentially identically to a T-junction of similar dimensions and wettability that is plasma bonded to glass or PDMS substrates.

A concern for any microfluidic device bonding technique is its biocompatibility. Hydrophobic surfaces, like those of our PDMS channels or of the adhesive of the tape, may adsorb biological molecules, depleting them from solution before they are encapsulated in drops, and interfering with downstream assays. To demonstrate that the method is sufficiently biocompatible so as to allow the encapsulation of commonly used biomolecules, including DNA and enzymes, we use our tape-bonded T-junction to form an emulsion for a droplet-based PCR. A PCR solution is prepared with 300 bp template DNA molecules, PCR primers, and Taq 1X Master Mix (New England BioLabs). The solution is divided in half, and one half is loaded into a PCR tube as the positive control and the other into a syringe. The solution in the syringe is then emulsified in fluorinated oil HFE-7500 with 4% (wt/wt) surfactant[47] using our tape-bonded drop maker and the

drops are collected into another PCR tube. Both tubes are thermocycled in a PCR machine and the emulsion is broken by adding a breaking solution of perfluorooctanol and HFE-7500 at a ratio of 40:60 by weight. The contents of the ruptured drops pool as an aqueous layer above the fluorinated oil and are removed with a pipette and visualized on an agarose gel, alongside the in-tube positive and negative (no-template) controls, as shown in Figure 3.4. The positive and droplet PCR both show distinct bands at the expected 300 bp amplicon length, while the negative control shows no such band. This illustrates that tape-bonded channels are compatible with droplet PCR.

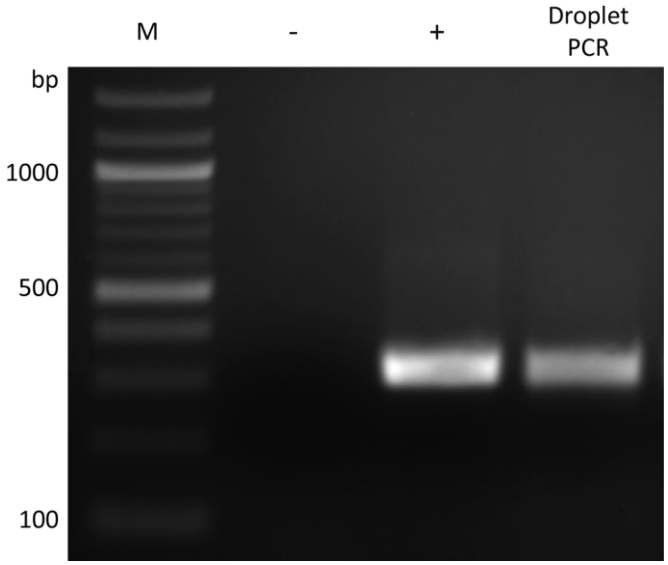


Figure 3.4 Droplet-based PCR using tape-bonded device. Ethidium bromide stained agarose gel used to visualize the amplicons produced from an emulsion PCR carried out with our tape-based drop maker. Distinct bands are visible for the droplet PCR and positive control. M: 100bp ladder; -: no template control; +: positive control

3.4 Conclusions

PDMS-tape bonding has several advantages over other bonding methods: It is inexpensive and uses materials that are commonly available. A variety of adhesive tapes can be used, including

transparent tapes that enable optical visualization of the channels. It allows devices to be bonded and used within 2 hours, which is convenient for rapid prototyping and testing. The bond is reversible, allowing the tape to be peeled away and the device washed to remove dust or contaminants, and then re-taped and re-used. Most importantly, because it does not require cleanroom facilities or a plasma bonder, it can be adopted by researchers who are not specialists in microfluidics. In addition, it should be useful for specialist microfluidic labs that simply want to bond and test their devices more quickly.

3.5 Supplemental Information

Adhesive tape surface roughness measurement

To measure the surface roughness of the adhesive coatings of our tapes, we first created molds of the tape surfaces. A sample of each tape was adhered to the bottom of a plastic Petri dish with the adhesive side facing upwards. PDMS elastomer (Sylgard) was prepared by mixing the elastomer base with crosslinker at a 10:1 weight ratio using a Dremel hand drill. The mixture was degassed under vacuum for 30 minutes, poured into the Petri dish, and cured for 2 hours at 65°C. After the PDMS was cured, thin cross-sections of each sample were cut using a razor blade. The surface of each cross-section was viewed using bright field microscopy with a 40x objective. ImageJ was used to quantify the surface roughness (R_a) for each sample.

Tape	Thickness	Roughness, R_a
Scotch® Magic™ Tape	.0635 mm	< 0.2 µm
Scotch® Permanent Double Sided Tape	.0762 mm	2.6 µm
Scotch® MultiTask Tape	.0584 mm	< 0.2 µm

Table 3.1. Adhesive tape specifications. Thickness values provided by the manufacturer.

Resolution of the Ra measurement is 0.2 µm.

Chapter 4: Sequence specific sorting of DNA molecules with FACS using 3DPCR

The following section is reprinted from the submitted manuscript, "Sequence specific sorting of DNA molecules with FACS using 3DPCR" by David J. Sukovich, Shea T. Lance, Adam R. Abate. David Sukovich, Shea Lance, and Adam Abate designed the research. David Sukovich and Shea Lance performed experiments. David Sukovich, Shea Lance, and Adam Abate wrote the manuscript.

In this chapter, we describe 3DPCR, which is a method for detection, quantification, and enrichment of specific DNA sequences out of a heterogeneous sample. 3DPCR encapsulates individual DNA molecules into double emulsion droplets and performs PCR reactions to identify droplets containing the DNA sequence of interest. The droplets containing the desired sequence are detected and sorted using Fluorescence-Activated Cell Sorting (FACS). There is an existing method for sequence detection and enrichment that requires a custom microfluidic sorter for molecule detection and enrichment[213]. Importantly, standard FACS machines are more widely available than custom microfluidic sorters, so 3DPCR is more readily accessible to researchers who do not specialize in microfluidics. By enriching highly heterogeneous samples for sequences of interest, 3DPCR allows for targeted deep sequencing such that reads are focused on DNA sequences of interest.

4.1 Abstract

Heterogeneity is an important feature of most biological systems, but introduces technical challenges to their characterization. Even with the best modern instruments, only a miniscule fraction of DNA molecules present in a sample can be read, and they are recovered in the form of short, hundred-base reads. In this paper, we introduce 3DPCR, a method to sort DNA molecules with sequence specificity. 3DPCR allows heterogeneous populations of DNA molecules to be

sorted to recover long target molecules for deep sequencing. It is valuable whenever a target sequence is rare in a mixed population, such as for characterizing mutations in heterogeneous cancer cell populations or identifying cells containing a specific genetic sequence or infected with a target virus.

4.2 Introduction

Heterogeneity is an important aspect of many biological systems, impacting their ability to respond to stimuli or evolve against evolutionary stresses. However, the massive heterogeneity present in most systems also imposes technical challenges to their characterization. A single ocean sample, for example, can contain billions of microbes, each with its own genes, gene clusters, and genomic structure[214]. Moreover, even the highest throughput modern sequencers are capable of sampling just a minute fraction of the sequences present in most systems, and the "reads" are obtained in the form of short, ~100 base fragments[215–217]. Genetic structures like genes, gene clusters, and genomes typically span much longer scales, from kilobases to gigabases of sequence[218–221]. To extract meaningful biological insight from the short read data requires "assembling" reads into "contigs" spanning the relevant genetic length scale. Here, however, the shortness of the reads, sparseness of the sampling, and massive diversity of the system impose technical challenges, making assembly difficult and error prone[217,222]. Often, useful contigs can only be generated for the most abundant organisms, and the overwhelming majority of species present at low and moderate levels are missed[223].

An effective means to overcome the challenges introduced by short read data is to enrich for target sequences prior to sequencing[224]. This can greatly improve assemblies and has been used to address questions ranging from deconvoluting cancer heterogeneity[225] to finding gene clusters encoding new, useful molecules in the microbiome[226]. A common enrichment strategy is PCR, in

which primers specifically amplify a target sequence out of a heterogeneous mixture. However, PCR methods can introduce artifacts[227,228], yield non-specific products[229], and are limited to enriching for sequences less than 10 kilobases in length. Digital droplet PCR (ddPCR) overcomes some of these issues and can be multiplexed to recover thousands of targets using microfluidics[213,230], but is still limited to lengths of a few kilobases[213,231,232]. An entirely different enrichment strategy is to physically purify molecules using beads coated in capture probes; however, this approach also is limited to targets of a few kilobases per capture probe, necessitating thousands of probes to recover megabases of sequence[233]. An optimal system for enriching DNA would significantly increase the lengths of target molecules that could be recovered per target probe, making enrichment more effective when little is known about the sequences of interest for target probe design.

In this paper, we introduce double emulsion digital droplet PCR (3DPCR), a method to physically sort long target molecules out of a heterogeneous population with FACS. Whereas PCR and oligo-capture methods are limited to a few kilobases of sequence length recovered, 3DPCR can theoretically exceed 100 kilobases. In addition, the PCR identification can be multiplexed, to recover molecules with multiple, distinct sequences – a new capability not possible with PCR or oligo methods that is useful for differentiating between organisms with combinatorial genetic diversity, such as microbes, viruses, and cancer cells.

4.3 Methods

Reagents and materials

Commercially available HFE, ionic and non-ionic (RAN Biotechnologies) Krytox-based surfactants were used to stabilize the double emulsions for thermal cycling. SU-8 was purchased from MicroChem. The Platinum Multiplex PCR Master Mix (PCR MM) containing dNTPs, PCR

buffer, and polymerase, as-well-as the Maxima SYBR Green/Rox qPCR Master Mix (2x) were obtained from ThermoFisher Scientific. Purified Lambda DNA and the ϕX174 Virion DNA were purchased from NewEngland BioLabs. Purified *Saccharomyces cerevisiae* DNA was purchased from Milipore. SYBR Green 1 Nucleic Acid Gel Stain was purchased from Lonza. PolyEthylene Glycol 35K (PEG35K), PolyEthylene Glycol 6K (PEG6K), Bovine Serum Albumin (BSA), perfluorooctanol (PFO), and Tween-20 were purchased from Sigma. Pluronic F-68 was purchased from Life-technologies. All oligonucleotides and probes used in this study (Table 4.1) were purchased from Integrated DNA Technologies.

Device Fabrication

The devices used to make double emulsions were fabricated in PDMS using soft lithography[201]. Planar masters composed of SU-8 were constructed using photolithography and used to mold PDMS devices. Inlet and outlet ports were punched using a 0.75 mm biopsy punch. PDMS devices were bonded to PDMS slabs by treating both with oxygen plasma for 60 s at 1 mbar of pressure in a plasma cleaner. Bonded devices were incubated at 65˚C for 36-48 hours prior to use[234].

SYBR Green PCR

Lambda DNA was diluted in DNAse-free water prior to use. For purification experiments, *S. cerevisiae* DNA was also utilized. PCR reagents were setup as such: 25 µL PCR MM, 4 µL Tween-20 (50%), 4 µL PEG 6K (50%), 1 µL primer EP04 (10 µM), 1 µL primer EP05 (10 µM), 1 µL diluted DNA, 14 µL DNAse-free water. Prior to use, the carrier phase inlet port and the emulsion outlet port of the bonded and baked devices were exposed to 1mbar of oxygen plasma

for 55 seconds[33]. Syringes containing (1) PCR mixture, (2) HFE oil supplemented with 2% PEG surfactant, and (3) carrier phase (10% PEG35K, 4% Tween-20, and 1% Pluronic F-68) were attached to their respective inlets via polyethylene tubing. Computer controlled syringe pumps were used to inject fluids at controlled volumetric flow rates (150 µL/hour for PCR mixture, 250 µL/hour for HFE supplemented with surfactant, and 1100 µL/hour for carrier phase) while being monitored visually on a microscope equipped with a short-shutter camera. Double emulsions were collected, supplemented with 1.5 mM $MgCl_2$, and cycled using a Bio-Labs thermocycler using the following parameters (86˚C for 2 min; 35x 86˚C for 30sec, 60˚C for 30 sec, 72˚C for 30 sec; 72˚C for 5 min). After thermocycling, samples were treated with 1x SYBR Green diluted in DMSO prior to visualization on an EVOS inverted fluorescence microscope.

TaqMan PCR

Lambda DNA and ϕX174 Virion DNA was diluted in DNAse-free water prior to use. PCR reagents were setup as followed: 25 µL PCR MM, 5 µL BSA (20%), 2.5 µL probe (5 µM), 0.5 µL primer (100 µM), 0.5 µL primer (100 µM), 2 µL diluted DNA, DNAse-free water to 50 µL. Prior to use, the carrier phase inlet port and the emulsion outlet port of the bonded and baked devices were exposed to 1mbar of oxygen plasma for 55 seconds[234]. Syringes containing (1) PCR mixture, (2) HFE oil supplemented with 2% ionic krytox, and (3) carrier phase (10% PEG35K, 4% Tween-20, and 1% Pluronic F-68) were attached to their respective inlets via polyethylene tubing. Computer controlled syringe pumps were used to inject fluids at controlled volumetric flow rates (150 µL/hour for PCR mixture, 250 µL/hour for HFE supplemented with surfactant, and 1100 µL/hour for carrier phase) while being monitored visually on a microscope equipped with a short-shutter camera. Double emulsions were collected, supplemented with 1.5 mM $MgCl_2$, and cycled

using a Bio-Labs thermocycler using the following parameters (86˚C for 2 min; 35x 86˚C for 30 sec, 60˚C for 90 sec, 72˚C for 20 sec; 72ºC for 5 min). After thermocycling, samples were visualized on an EVOS inverted fluorescence microscope.

Flow cytometry of 3DPCR reactors

Double emulsions were commonly sorted using the FACSAriaII system from BD. Samples were diluted in a diluent (2% Pluronic F-68, 1% PEG35K) and loaded onto the FACSAriaII via the stage. Double emulsions were first gated using SSC and FSC diffraction, followed by gating using the presence or absence of the SYBR, FAM, or Cy5 fluorophore. SYBR Green and FAM fluorescence were identified using a 488 nm laser and a 505LP optical filter (BD Biosciences). The Cy5 fluorophore was identified using a 633 nm laser with a 670/30 filter.

Quantitative PCR

Quantitative (q)PCR was performed using a Stratagene Mx3005P machine. Prior to running, DNA input was normalized to the number of emulsions collected. For identification of (lambda) targets pre- and post-sorting, primers EP07 and EP08 were used in the qPCR reaction. For indentification of offtarget (*S. cerevisiae*) DNA presence pre- and postsorting, Ye01 and Ye02 were used in the qPCR reaction. In general, cycling conditions used were 95˚C for 5 min, followed by 40x cycles of 95˚C 30sec, 55˚C 1 min, and 72˚C 30 sec.

4.4 Results

3DPCR is a cousin of common ddPCR, except that rather than performing single-molecule PCR assays in water-in-oil single emulsion droplets, it performs the reactions in double emulsion droplets. The benefit of this is that double emulsions, unlike single emulsions, are suspended in an

aqueous carrier phase, allowing them to be read and sorted using common flow cytometry instruments[235]. Using specific PCR reactions performed in each droplet, each molecule in the sample is "read" at a specific location to determine if it is a target molecule. If so, the PCR assay yields a fluorescent signal that fills the encapsulating droplet, allowing it to be detected and recovered by FACS sorting. Using TaqMan PCR with probes of different color, the method can be multiplexed, to differentiate between molecules that share partial homology and sort based on combat atrial rules, such as a given molecule containing two specific sequences.

The objective of 3DPCR is to enable the detection, quantitation, and isolation of specific DNA molecules in a heterogeneous sample. To accomplish this, 3DPCR encapsulates DNA from a sample into double emulsion droplets (Figure 4.1A) and performs a specific PCR reaction in every droplet interrogating for target sequences (Figure 4.1B). If the sequences are present, PCR amplification occurs, generating a fluorescent signal that can be detected on FACS (Figure 4.1C) and used to sort and recover the target sequences (Figure 4.1D). Sorted droplets can be pooled and analyzed as a mixture, or dispensed singly into wells. While we use standard PCR techniques for detecting amplification, with SYBR or TaqMan fluorescence as the readout, other reactions and readouts are applicable, including isothermal amplification (MDA, LAMP, RPA) and probes (molecular beacons, scorpion probes).

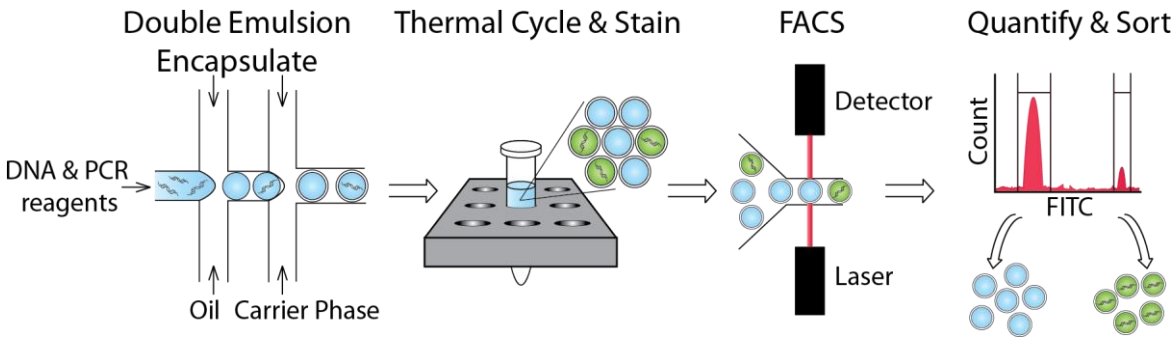


Figure 4.1 Overview of 3DPCR workflow. A heterogeneous sample of DNA is encapsulated into double emulsion droplets microfluidically (A) and the droplets are thermal cycled (B). Droplets contain the target sequence undergo amplification, producing a fluorescent signal that can be detected on FACS (C) and used to quantitate and sort the droplets (D).

3DPCR double emulsions

3DPCR is made possible by the ability of microfluidic devices to form monodisperse droplets from a sample containing a mixture of DNA molecules. By controlling the concentration of molecules in the sample, the number encapsulated in each droplet can be tuned down to single molecules[236]. Critical to using double emulsions for PCR detection of target molecules is stabilizing the droplets through thermocycling reactions. In this work we accomplish this using fluorinated oils with fluorosurfactants due to the incredible stability of the emulsions they form, including for the high temperatures of PCR. We use two surfactants depending on the PCR assay conducted in the droplets: a non-ionic polyethylene glycol (PEG) Krytox surfactant, and an ionic-Krytox surfactant (Figure 4.2).

The non-ionic surfactant comprises a PEG head group covalently bound to two fluorinated block copolymer tails; the PEG is soluble in water, while the fluorinated tails are soluble in the oil phase of our emulsions[47]. When this amphiphilic molecule adsorbs to the water-oil interface of a droplet, the PEG head group coats the inner-surface of the water droplet. PEG is effective at preventing the adsorption of proteins to hydrophobic interfaces, which is important for maintaining the enzymes in the bulk of the droplet, where they can carry out PCR. To detect the amplification products of the reactions, we add SYBR green dye to the carrier phase after amplification. SYBR green becomes fluorescent when intercalated into DNA[237] and readily partitions through the thin

oil shell of the double emulsions, staining the positive droplets and making them detectable via FACS. The double emulsions formed with this surfactant are monodisperse and survive PCR thermocycling (Figure 4.2A).

While the non-ionic surfactant provides good stability and enables efficient PCR in double emulsions, the shells are leaky to dye molecules; indeed, we exploit this to stain positive droplets with SYBR post-thermocycling. However, this leakiness also creates a challenge when performing TaqMan PCR in the droplets: the reaction uses DNA probes labeled with a fluorescent dye and quencher; upon amplification, the probe is cleaved releasing the dye to fluoresce. However, the untethered dye can readily partition through the double emulsion shell, leaking out of the droplets and resulting in loss of the TaqMan signal.

To rescue the signal, we must contain the dye, which we achieve by introducing a barrier to its partitioning through the shell. We switch to the ionic-Krytox surfactant which, in contrast to the PEG surfactant, readily adsorbs proteins to the droplet interface[47]. This would normally be detrimental to PCR, as it would adsorb the polymerases, but we also include bovine serum albumin (BSA) in the droplet at high concentration. BSA is often added to PCR to increase efficiency of the reaction and, being a protein, is readily adsorbed to the interface of ionic-Krytox stabilized droplets. At the interface, it forms a "skin" that prevents further adsorption of proteins and provides a barrier to partitioning of the freed dye molecules. This makes TaqMan positive droplets detectable on FACS. This surfactant also forms monodisperse double emulsion droplets stable to thermocycling (Figure 4.2B).

The double emulsion droplets, regardless of the surfactant used, swell during PCR (Figure 4.2C). We believe this to be due to partitioning of water and buffer molecules through the shells, to balance the osmolarities of the inner and outer aqueous phases. Shell permeability must thus be

carefully considered when performing reactions in double emulsion droplets, but also provides a facile means of modulating droplet contents by adding compounds to the carrier phase.

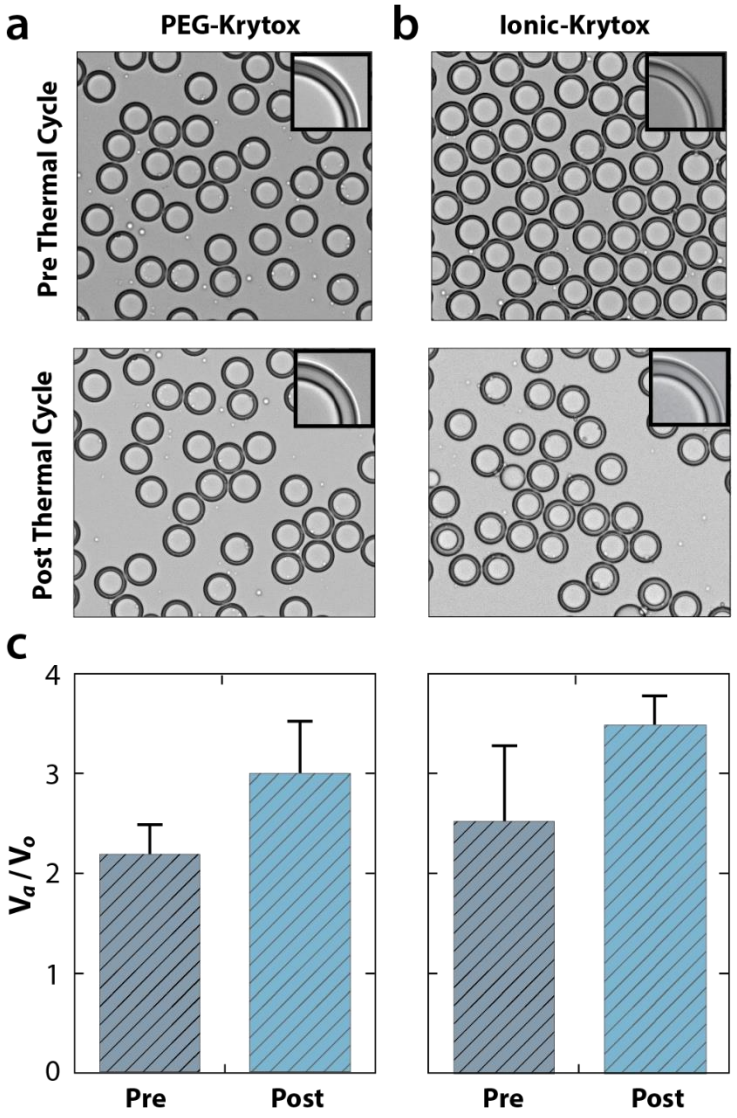


Figure 4.2 Double emulsion monodispersity. We form double emulsions microfluidically stabilized by either a non-ionic PEG-Krytox (A) or ionic-Krytox surfactant (B). Both yield extremely stable droplets that survive thermal cycling. However the droplet swells during thermal cycling to balance the osmolality of the inner core of the droplet with the miscible aqueous carrier.

One-color 3dPCR with SYBR readout

3DPCR, like ddPCR, can be used to count target DNA molecules in a heterogeneous mixture. To demonstrate this, we analyze samples of Lambda virus at different concentrations (Figure 4.3). We spike lambda DNA into PCR mix with primers targeting a portion of the Lambda genome (35,515-35,664bp). We double-emulsify and thermocycle the sample, then add SYBR to the carrier phase to stain the positive droplets (Figure 4.3A). To analyze the ~5 million double emulsion droplets, we use a flow cytometer. The double emulsions are monodisperse and large compared to cells (~40 μm) and, thus, appear as a tight cloud at high intensity in the forward and side scatter channels (Figure 4.3B, *left*). Debris appears at lower scatter values and oil droplets resulting from ruptured double emulsions as a distribution of low side- and high forward- scatter events. Large, multicore double emulsions are occasionally produced by the microfluidic drop maker and appear as a relatively narrow and very high side- and forward-scatter population. To analyze the single-core double emulsion population, we gate the corresponding cloud in the scatter channels (red circle), and plot the fluorescence of this subpopulation (Figure 4.3B, *right*). We observe two distinct populations in the fluorescence channel, a dim one corresponding to negative droplets devoid of the target, and a bright one corresponding to droplets containing the target. To measure the target concentration in the sample, we measure the proportion of dim and bright droplets, normalizing by volume to convert to concentration units. We perform this analysis on six samples varying in Lambda virus concentration by ~5 orders-of-magnitude (Figure 4.3C) and find that, as expected, the fraction of bright droplets scales in proportion to the input concentration.

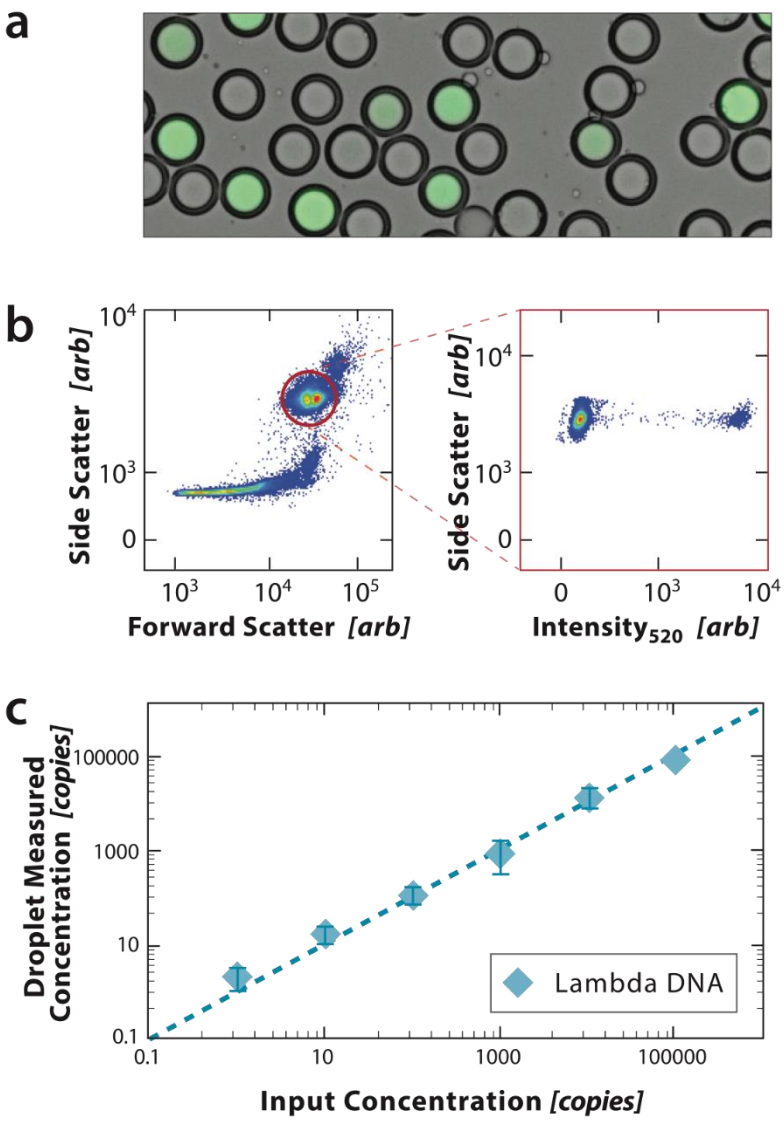


Figure 4.3 3DPCR with a one-color SYBR assay. Lambda virus DNA is mixed with primers targeting the virus and PCR reagents, formed into double emulsions, thermally cycled, and stained with SYBR green (A). The droplets are processed through FACS, gated on scattering to discard all non-single-core double emulsion events, the remainder for which are plotted the fluorescence values (B). Six samples with different lambda virus concentrations are processed and quantified, demonstrating that, as expected, the proportion of fluorescent droplets scales with the lambda virus input concentration, in accordance with Poisson encapsulation statistics.

Multiplexed 3DPCR with a TaqMan readout

TaqMan PCR increases the specificity of conventional PCR by using a specific TaqMan probe to generate the fluorescence readout. Only if the amplification products are homologous to the TaqMan probe sequence will the fluorescent signal be generated, reducing false positives due to non-specific amplification[238]. TaqMan PCR also allows the reaction to be multiplexed by using dyes of different color to label the probes, to interrogate for multiple DNA sequences in the sample simultaneously. This would allow us, for example, to distinguish between droplets containing a single sequence and ones containing multiple sequences. To illustrate this, we spike Lambda virus genomes into a sample of ΦX174 virus genomes, varying the concentration of Lambda virus over ~5 orders-of-magnitude while holding that of ΦX174 fixed. We analyze the sample with 3DPCR using TaqMan assays targeting Lambda (green dye) and ΦX174 (red dye) virus (Figure 4.4A). The Lambda and ΦX174genomes do not associate with one another and are encapsulated randomly in the droplets; we thus expect four populations of droplets, ones devoid of both targets (dim, [- , -]), ones with one target (green [+ , -] or red [- , +]), and ones with both targets (yellow, [+ , +].

We analyze the signal with FACS and again gate the scatter population corresponding to the double emulsions (Figure 4.4B, *left*), plotting the red and green fluorescence of this subpopulation and observe the expected four fluorescence populations (Figure 4.4B, *right*). Just as with the one-color experiment, the proportion of droplets falling within the different fluorescence populations can be used to estimate the concentrations of the different viral genomes in the mixed sample over a wide dynamic range (Figure 4.4C). This shows that 3DPCR with a TaqMan readout can quantitate multiple DNA species in a sample simultaneously. Moreover, it allows specific instances in which two different targets are present within the same droplet to be identified (yellow droplets, Figure 4.4A). This is not possible with conventional qPCR and is

useful for correlating different sequences together, for example, to characterize the lengths of DNA molecules based on the presence of sequences at defined distances from one another[239,240] to characterize recombination events based on the presence of different sequence combinations within a single molecule[241], and to characterize combinatorial associations of distinct molecules within the same entity, such as the DNA of viruses with segmented genomes[242,243] or the probability of a specific pathogen to infect a given host. Such associations, normally, are only measurable with DNA sequencing, but 3DPCR provides a rapid and inexpensive alternative.

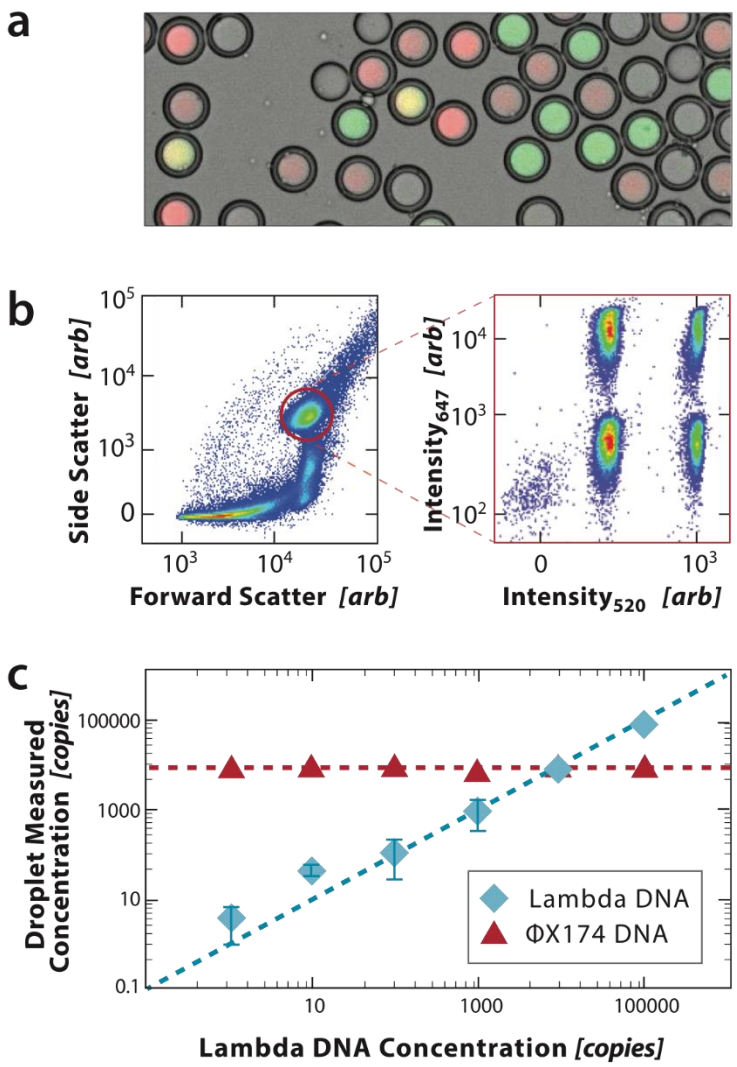


Figure 4.4 3DPCR with a two-color TaqMan assay. Lambda- and ϕX174 virus DNA are mixed together and combined with primers and TaqMan probes targeting both viruses. The samples are formed into double emulsions and thermally cycled (A). The droplets are processed through FACS, gated on scattering to discard all non-single-core double emulsion events, the remainder for which are plotted the fluorescence values (B). Four fluorescent populations are observed, corresponding to the four possible combinations of lambda and ϕX174 virus encapsulation. Six samples with different lambda virus concentrations with constant ϕX174 virus conditions are processed, and quantified, demonstrating that, as expected, the proportion ϕX174 positive droplets is unchanged

between samples, but that of Lambda virus scales with the input concentration, in accordance with Poisson encapsulation statistics (Dashed curve).

Sorting single DNA molecules with FACS

3DPCR allows a FACS instrument to "read" a target DNA sequence using a PCR assay. However, FACS is capable of more than detection, it can also sort based on the measurement. This, in essence, allows a mixed population of DNA molecules to be interrogated individually, to recover all molecules matching the PCR assay criterion; it also provides a novel way to enrich DNA that affords substantial advantages over conventional methods based on common PCR or oligo capture. To demonstrate this ability, we sort DNA samples using 3DPCR. We spike Lambda virus DNA into a background of *S. cerevisiae* genomic DNA, and sort the sample with TaqMan probes targeting Lambda (Figure 4.5A, *left*). The sorted double emulsion droplets are TaqMan positive, although there are also a significant number of oil droplets (Figure 4.5A, *right*). The oil droplets are the remnants of double emulsions that explode during FACS detection, ruptured by the high shear rates generated by the sheath flow of the instrument. Such rupture can be mitigated by varying double emulsion size and the nozzle size of the FACS.

To confirm the enrichment achieved by sorting the double emulsions, we use qPCR, analyzing the sorted and unsorted samples with primers targeting a different region of the Lambda genome and a region of the *S. cerevisiae* genome (Figure 4.5B). The Lambda virus primers amplify slightly sooner in the sorted than the unsorted sample (Figure 4.5B, *left*), even though the total amount of DNA is the same, indicating that sorting increases the number of Lambda virus genomes. Conversely, the *S. cerevisiae* primers amplify much later in the sorted than in the unsorted sample, indicating that *S. cerevisiae* DNA has been strongly de-enriched by sorting

(Figure 4.5B, *right*). Based on the qPCR curve shifts, we estimate the ratio of Lambda- to *S. cerevisiae* DNA has changed by a factor of 83; this is close to the theoretical expectation based on Poisson encapsulation of the molecules for our loading rate of 1% Lambda-positive droplets. Higher enrichment ratios can be achieved by diluting the sample further, reducing the amount of *S. cerevisiae* DNA encapsulated in the positive droplets; however, this also reduces the frequency of positive droplets, necessitating more sorting to recover the same number of positive events. This demonstrates that 3DPCR is an effective means of sorting target DNA out of a mixed sample with FACS.

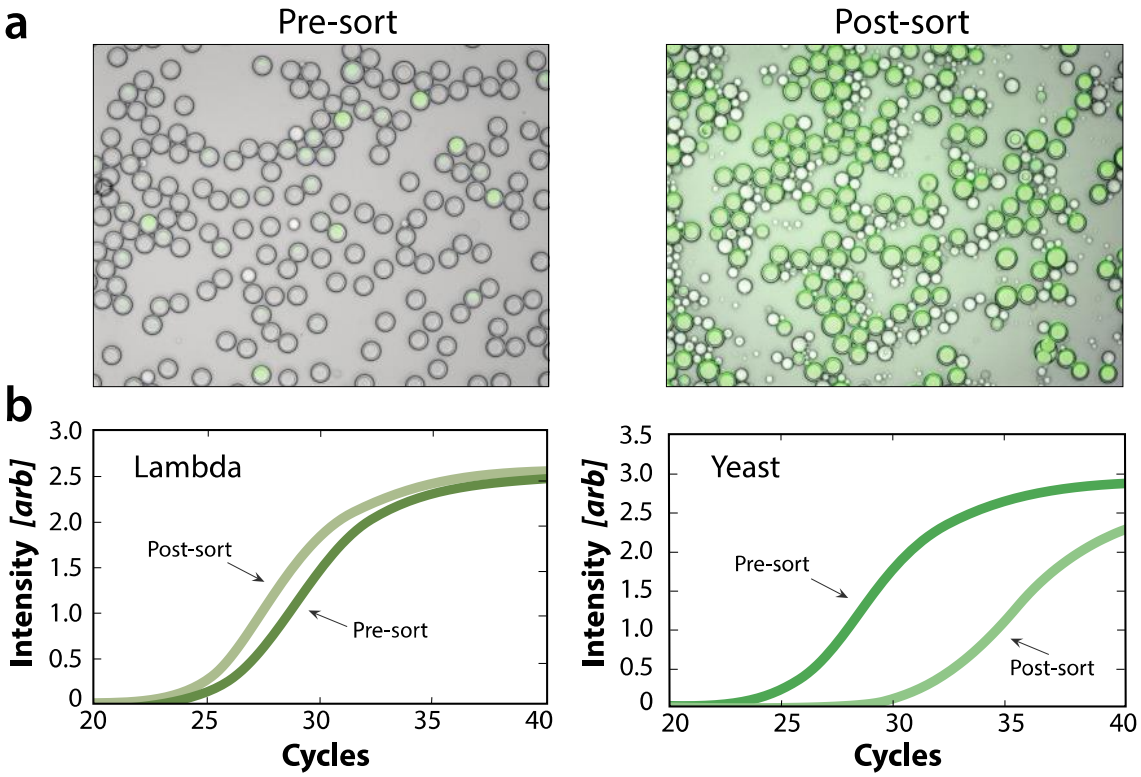


Figure 4.5 Enrichment of Lambda virus DNA out of a background of S. cerevisiae (Yeast) genomic DNA by sorting the molecules with FACS. Lambda DNA is spiked into S. cerevisiae genomic DNA at 1:100 concentration, and processed through (A) 3DPCR (left) and FACS (right). Some double emulsions pop during FACS, leaving behind small oil droplets in the sorted

population. We quantify enrichment using qPCR with primers targeting a region of the S. cerevisiae genome and a different region of the Lambda genome that was amplified in the 3DPCR reaction. Based on the curve shifts, we estimate that the ratio of Lambda-to-S. cerevisiae DNA increases by 83 fold by sorting.

4.5 Conclusions

3DPCR allows DNA molecules to be read and sorted by FACS, providing a novel and powerful means of enriching target molecules. Because 3DPCR can be multiplexed using TaqMan probes of different color, it also allows for characterization of associations between distinct sequences, such as sequences existing within the same molecule, virus, or cell. While we limit our analysis to DNA, RNA molecules can also be sorted with 3DPCR, only requiring that an additional step of reverse transcription be added, similar to as has been reported previously for ddPCR. Cells and viruses may also be sorted with the method, although this introduces the technical challenge of lysing these entities, so that their nucleic acids are available for the PCR analysis. While detergents present in PCR buffers and the high temperatures required for the reaction may be sufficient to lyse many organisms, it may not be for others, such as thermo-tolerant viruses. 3DPCR adds new and powerful capabilities to digital PCR and makes the approach useful for enriching DNA for the first time. It is useful for studying a host of complex biological systems exhibiting heterogeneity, such as the cell and virus genomes that comprise organismal tissues and tumors, and microbial ecologies inhabiting the human gut and nearly every habitat on the planet.

4.6 Supplemental Information

Name	**Sequence (5'-3')**
EP01	GACTACATCCGTGAGGTGAATG
EP02	TTGCGCTGCTTATGCTCTAT
EP03	/5Cy5/TCTGCCCGTGTCGGTTATTCCAAA/3IAbRQSp/
EP04	CTTTGAATGCTGCCCTTCTTC
EP05	CAGATAACCATCTGCGGTGATA
EP06	/56-FAM/ATACTGAGC/ZEN/ACATCAGCAGGACGC/3IABkFQ/
EP07	CGGTAGAGGTAAAGGTGGTAAAG
EP08	CTGATAGCTGGCTTAGTGATACC
PX01	GCGCTCTAATCTCTGGGCAT
PX02	CAAAGAAACGCGGCACAGAA
PX03	/56-FAM/ATGGAACTG/ZEN/ACCAAACGTCGTTAGGCCA/3IABkFQ/
Ye01	CGGTAGAGGTAAAGGTGGTAAAG
Ye02	CTGATAGCTGGCTTAGTGATACC

Table 4.1 Primers and TaqMan probe sequences

Chapter 5: Robust and specific sorting of viruses with PCR-Activated Virus Sorting (PAVS)

The following section is reprinted from the submitted manuscript, "Peering below the diffraction limit: Robust and specific sorting of viruses with flow cytometry" by Shea T. Lance, David J. Sukovich, Kenneth M. Stedman, and Adam R. Abate. Shea Lance, David Sukovich, Kenneth Stedman, and Adam Abate designed the research. Shea Lance and David Sukovich performed experiments. Shea Lance, Kenneth M. Stedman, and Adam Abate wrote the manuscript.

In this chapter, we describe PCR-Activated Virus Sorting (PAVS), which is a method for detection, quantification, and enrichment of specific viruses out of a heterogeneous sample. PAVS uses the 3DPCR technique described in Chapter 4 to encapsulate individual viruses into double emulsion droplets and perform PCR reactions to identify droplets containing the virus of interest. The droplets containing the desired virus are detected and sorted using Fluorescence-Activated Cell Sorting (FACS).

Studying viruses using conventional methods is challenging because their small size precludes direct sequencing of individual virions or visualization using optical microscopy. Additionally, the majority of viruses cannot be cultivated. PAVS is a powerful technique for virological study, as it enables researchers to detect and enrich for rare or and uncultivable viruses We show that full-length viral genomes can be detected using multiplexed PCR, which is useful for characterizing the integrity of viral genomes. We also demonstrate how PAVS can used for targeted recovery of single virions can be from a mixed sample, which can then be subjected to downstream analyses such as qPCR.

5.1 Abstract

Viruses are incredibly diverse organisms and impact all forms of life on Earth; however,

individual virions are challenging to study due to their small size and mass, precluding almost all direct imaging or molecular analysis. Moreover, like microbes, the overwhelming majority of viruses cannot be cultured, impeding isolation, replication, and study of interesting new species. Here, we introduce PCR-activated virus sorting, a method to isolate specific viruses out of a heterogeneous population. Specific sorting opens new avenues in the study of uncultivable viruses, including recovering the full genomes of viruses based on genetic fragments in metagenomes, or identifying the hosts of viruses.

5.2 Introduction

Viruses impact every form of life on earth, from applying evolutionary stresses to enhancing the transfer of genes between organisms[244–247]. Many human diseases are caused by viruses, including acute diseases like Ebola[248] and influenza[249], and chronic diseases caused by Epstein-Barr Virus[250], Human Immunodeficiency Virus[251], and recently discovered Zika virus[252]. Studying viruses is thus important to human health, but also for elucidating the incredible mechanisms they've evolved to survive, replicate, and spread; these discoveries may lead to new molecular techniques and methods for treating disease. Studying viruses, however, can be challenging. They are usually much smaller than the diffraction limit of light and thus not directly visible with optical microscopy. They contain miniscule amounts of nucleic acid and protein, making direct sequencing or proteomic characterization of individual virus particles challenging[253]. To overcome these issues, the standard strategy is to culture the virus of interest to produce sufficient quantities for biological assays, such as gel electrophoresis, infection assays, or visualization with super-resolution or electron microscopy. However, like microbes, most viruses cannot be cultured[254], as this requires knowledge of which host cells the virus replicates in which, for most viruses, are also likely uncultivable[255,256].

When a virus cannot be cultured, molecular methods are valuable. For example, viruses can be purified from a sample using filtration or density-dependent centrifugation, to recover particles of the appropriate size range, and the nucleic acids purified for PCR or next generation sequencing[257,258]. This can be applied directly to environmental viruses and provides a genetic snapshot of organisms in that environment, and has yielded numerous insights into virus phylogeny and fundamental biology[257,258]. However, viruses are also the most diverse organisms on the planet and viral samples often comprise sequences from trillions of entities, exceeding by orders of magnitude the limits of modern sequencers to sequence them[259]. As a result, such "shotgun" sequencing provides a sparse sampling of the system recovered as billions of short, hundred-base reads[216]. To extract meaningful biological insight from this complex data, the reads must be pieced into viral genomes, introducing substantial bioinformatic challenges that, often, cannot be overcome[215,253,260]. Most often, only genomic sequences for the most abundant organisms can be recovered and little is learned about the vast number of new viruses present at low-to-moderate levels[223,261]. To enhance the investigation of viral ecosystems, a method for culture-free purification of specific species would be valuable; however, as of yet, no method exists for specific sorting of viruses.

In this paper, we present specific and high throughput sorting of viruses, PCR-Activated Virus Sorting (PAVS). Using microfluidics, we encapsulate single particles of a population of diverse viruses into monodisperse double emulsion droplets. PCR reagents targeting specific genetic loci are also included, interrogating every droplet for these sequences. If a virus contains them, PCR signals are generated that cause the droplet to become fluorescent, making it sortable by double emulsion flow cytometry[235]. The recovered droplets can be ruptured and the material subjected to additional analyses, such as quantitative PCR or digital PCR and sequencing. The

approach is simpler than antibody-based labeling and sorting of cells[262] because designing PCR assays for specific detection of sequences is much easier than generating high affinity antibodies with which to label and sort single virus particles by flow cytometry. Moreover, the implementation of TaqMan PCR allows multiplexing to interrogate each virus for distinct sequences. As we show, multiplexing can be used to measure the length distributions of viral genomes in a sample and is extendable to sequences that are not physically connected, such as genomic segments of viruses like influenza, or the 16S rRNA sequence of a bacterial cell harboring the target virus. Flow cytometric sorting has become a universal tool in cell biology and microbiology and PAVS allows it to be applied to viruses for the first time.

5.3 Materials and Methods

Preparation of viral samples

Bacteriophage T4 (T4) and bacteriophage ΦX174 (ΦX174) (from Carolina Biological Supply) is propagated by infection of *E.coli* B (ATCC 11303) and ΦX174 by infection of *E.coli* C (ATCC 13706), respectively. Bacteriophage lambda cI857ts is obtained from the lambda lysogen *E.coli* strain KL470 and propagated by infection of *E.coli* C600.

Microfabrication of devices

The microfluidic devices are fabricated using soft lithography in poly(dimethylsiloxane) (PDMS)[201]. SU-8 masters are fabricated by photolithography and used to mold PDMS devices by mixing PDMS polymer and cross-linker at a ratio of 11:1, pouring over the master, degassing to remove air bubbles, and baking at 75ºC for 4 hours to solidify. The device is extracted from the master with a scalpel, and inlet and outlet ports added with a 0.75 mm biopsy punch (Harris, Unicore). The device is washed with isopropyl alcohol and patted with scotch tape to remove

debris prior to plasma bonding. The flow focus drop maker is bonded to a glass slide, baked at 75ºC for 15 minutes, and treated with Aquapel to render the channels hydrophobic for water-in-oil emulsification. The double emulsion device is bonded and baked at 75ºC for 48 hours to completely revert the wettability to its native hydrophobic state. To pattern the channel wettability for double emulsification, select ports are blocked with Scotch tape, leaving others open for oxygen plasma treatment of 1 minute[234].

Encapsulation of viruses in double emulsion droplets

The virus samples are mixed with PCR reagents (Platinum Multiplex PCR Master Mix, Thermo Fisher) and PCR primers (IDT) specific for the species of interest. The middle phase of the double emulsion consists of HFE-7500 fluorinated oil (3M) with 2% (w/w) PEG-PFPE amphiphilic block copolymer surfactant[47], and the carrier aqueous phase of 4% (v/v) Tween 20, 1% (v/v) Pluronic F-68 (Gibco), and 10% (w/v) PEG (molecular weight 35 KDa) in water[234]. These solutions are loaded into syringes (BD 1 mL luer lock; 27G ½" needle), and the virus and PCR solution into a syringe atop 200 µL HFE-7500 oil; the oil acts as a hydraulic to push the solution into the device to accommodate for dead volumes, allowing nearly all of the solution to be used. The syringes are mounted onto pumps (New Era) with needles (BD), polyethylene tubing (PE-2) is affixed to the needles, and the syringes are primed by flowing at 5,000 µL/h prior to connecting them to the device. Flow rates are controlled with a custom Python script and set to 90 µL/h for the virus sample, 80 µL/h for the oil, and 250 µL/h for the outer aqueous phase. The double emulsion droplets exit the device through PE-2 tubing and are collected into a 1.5 mL microcentrifuge tube.

To prepare the double emulsion droplets for thermal cycling, the sample is transferred from the 1.5 mL microcentrifuge tube into 0.2 mL PCR tubes, such that each contains 90 µL of emulsion and 10 µL of fresh PCR buffer; the PCR buffer consists of 30 µL of 50 mM $MgCl_2$ and 100 µL of 200 mM Tris pH 8.0 and 500 mM KCl and is essential for ensuring that PCR components do not leach out of the droplets into the carrier phase, in which they are soluble. The sample is cycled on a T100 thermal cycler (Bio-Rad) according to the Platinum Multiplex Master Mix instructions. After thermal cycling, 1x SYBR Green I (Life Technologies) is loaded into the carrier phase, permeating through the double emulsion shell and staining the droplets that have undergone PCR amplification.

FACS detection and sorting of droplets

A fluorescence-activated cell sorter (FACS) Aria II (BD) is used to sort the emulsions to recover droplets that contain the virus of interest. The FACS chamber temperature is set to 4°C and agitation speed to the highest setting to prevent the droplets from sedimenting during the sort. The droplets strongly scatter the FACS laser, requiring a 2x Neutral Density (ND) filter to decrease signal into the detectable range. The microfluidic device produces uniform double emulsions and, consequently, the droplets appear as a compact cluster in forward versus side scatter, making them easy to distinguish from particulate and small oil droplets, which are ignored in the analysis.

The sample is analyzed in batches by diluting 100 µL of emulsion into 200 µL of 2% (v/v) Pluronic F-68 and 1% (w/v) PEG (molecular weight 35 K) in water, and gently mixing using a 200 µL pipette tip. The sample is loaded into the FACS and the double emulsions gated in the Forward Scatter (FSC) and Side Scatter (SSC) channels[235]. To read the SYBR channel relating to amplification, we use 488 nm excitation with 520 nm emission; the population has two peaks, one

with low average intensity representing empty or negative droplets, and another with high average intensity representing SYBR positive droplets, which we gate to recover. We use the strict "purity" setting of the instrument which discards events in which multiple droplets pass through the detection window at the same time.

Amplification of recovered viral DNA

Sorted droplets are briefly centrifuged to localize them at the bottom of the tube. To release nucleic acids, the droplets are ruptured by adding 20 µL of DI water and 50 µL of perfluorooctanol (PFO), and vortexing for 1 minute. The sample is centrifuged again, and the aqueous top phase containing the viral DNA removed using a micropipette.

To confirm enrichment of T4 phage in the sorted emulsion, we use quantitative PCR (qPCR). The concentration of viral DNA is too low post-sorting to be reliably detected by bulk qPCR. To address this, we non-specifically amplify the material using digital droplet multiple displacement amplification (ddMDA) prior to qPCR analysis using the Qiagen REPLI-g Single Cell Kit. ddMDA is a non-specific method that amplifies all nucleic acids in a sample uniformly[263]. The sample is incubated with 3 µL of the Denaturation Solution for 10 minutes at 65°C. After heating, the reaction is halted by adding 3 µL of Stop Solution Mix. 20 µL of the REPLI-g sc Master Mix containing 14.5 µL of REPLI-g sc Reaction Buffer, 4.5 µL of water, and 1 µL of REPLI-g sc Polymerase is added to each sample. The sample is encapsulated into droplets using a 20 µm flow-focus drop maker[54,63] and HFE-7500 fluorinated oil with 2% (w/w) PEG-PFPE amphiphilic block copolymer surfactant. The emulsion is collected into an 1.5 mL microcentrifuge tube and the reaction incubated at 30°C for 16 hours. After incubation, the droplets are ruptured by adding 10 µL of PFO, vortexing, and spinning as above.

Quantitative PCR analysis of sorted droplets

To confirm that PAVS enriches for bacteriophage T4 over bacteriophage ΦX174, we estimate the concentrations of both viruses in the sorted and unsorted pools using qPCR (Stratagene Mx3005P, Agilent). The qPCR primer sequences are different from the ones for PAVS detection, so that in-droplet amplification products do not skew the qPCR results. Cross-threshold (C_t) values for T4 and ΦX174 in the pre- and post-sorted samples are used to compute an enrichment factor. The amplification reagent for all the qPCR measurements is Maxima SYBR Green Master Mix (Thermo Scientific), and the qPCR primers are listed in Table 5.1.

5.4 Results

PAVS enriches for specific viral species from a heterogeneous sample

PCR-Activated Virus Sorting allows specific viruses in a mixed population to be detected and recovered by sorting. This is accomplished by encapsulating the viruses in double emulsion droplets using microfluidic technology and performing PCR in each droplet to probe for sequences of interest. Because the viruses are encapsulated at 0.1 per droplet, most droplets are empty or contain a single virus, in accordance with Poisson statistics (Figure 5.1). If the target virus is present in a droplet, PCR amplification occurs, generating a fluorescent signal that can be detected and recovered by FACS. Due to the rapid rate at which microfluidics can encapsulate individual viruses in droplets (>1 KHz), millions of single virus particles can be sorted in a few hours.

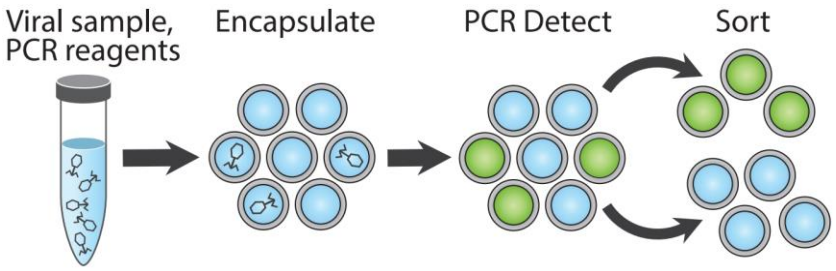


Figure 5.1 PCR-Activated Virus Sorting (PAVS) workflow. Virus suspensions are first encapsulated with PCR reagent and probe in double emulsion droplets, then thermal cycled and stained with SYBR green. FACS is used to sort fluorescent droplets containing the viral species of interest, generating an enriched sample that is ready for downstream processing.

Specific detection and quantification of viral genomes

To identify a virus in a droplet, PAVS uses a PCR assay interrogating for sequences that exist within the target species. In this way, the PCR primers are analogous to antibodies when sorting cells with FACS, providing a detectable fluorescence signal only when the target species is present. However, whereas generating high affinity antibodies against a target virus can be challenging, especially if it is uncultivable, designing specific PCR primers is straightforward. This makes PAVS general, allowing it to recover any virus of interest to which PCR primers can be designed. To illustrate this, we perform digital TaqMan PCR on samples containing T4, ΦX174, and lambda virus, using probes specific for only bacteriophage T4 (**S1**). After thermal cycling, we observe TaqMan positive droplets in the T4 sample, demonstrating successful amplification when this virus is present, as expected (Figure 5.2A). By contrast, TaqMan fluorescence is absent in the ΦX174 and lambda negative controls (Figure 5.2A), confirming that the reaction is specific. This shows that a single virus TaqMan PCR assay can be used to differentiate between these viruses.

In addition to enabling the detection of specific viruses in a sample, PAVS can count individual virus particles. To demonstrate this, we analyze a dilution series of T4 bacteriophage, reading out the results with fluorescence microscopy and image analysis (ImageJ). We find that, as expected, the fraction of positive droplets is directly proportional to T4 concentration (Figure 5.2B). This is due to the viruses being loaded at limiting dilution, such that most droplets are empty but a small fraction contains virus particles. Under such conditions, the viruses are encapsulated individually and the number of droplets containing a virus is approximately equal to the number of viruses in the sample, in accordance with random Poisson encapsulation.

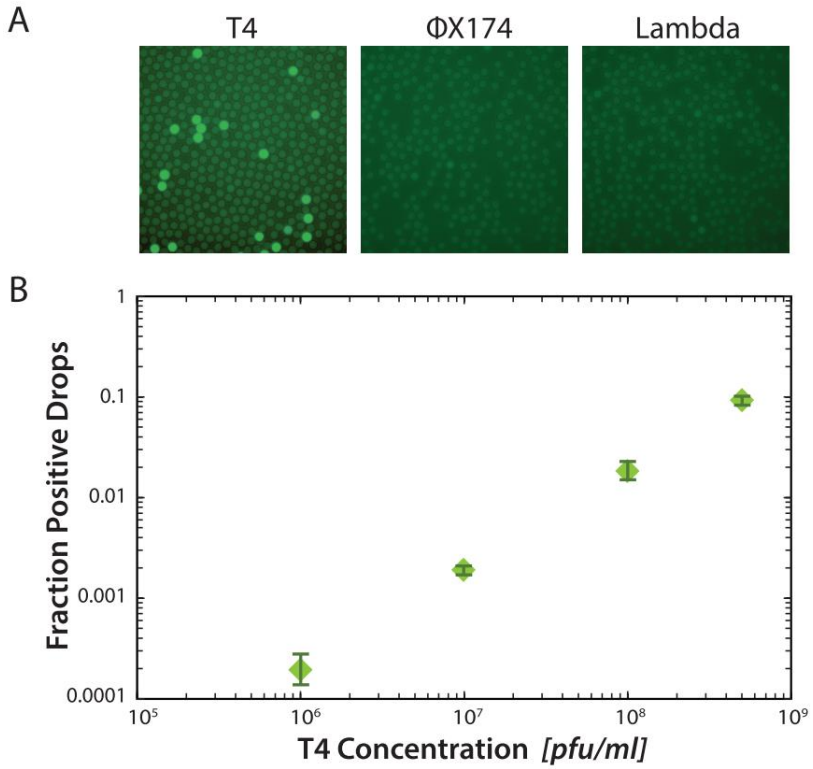


Figure 5.2 Specific detection and quantitation of viral genomes. A) T4, ΦX174, and lambda virions are partitioned into droplets with TaqMan primers and probe specific for T4. After thermal cycling, the T4 sample has TaqMan signal while the ΦX174 and lambda negative controls have no signal, demonstrating that digital droplet PCR specifically detects target viruses. B) The fraction of TaqMan positive droplets in digital PCR for T4 is closely related to the input T4 concentration,

showing that digital droplet PCR quantitatively measures viral concentration. Error bars indicate the standard deviation for triplicate measurements.

Multiplexed digital PCR can detect full-length virus genomes

A unique and valuable property of PAVS is that it can differentiate between viruses that contain just one target sequence and others that contain multiple. This is possible because TaqMan PCR can be multiplexed using probes targeting different sequences labeled with fluorescent dyes of different color. Hence, viruses containing one sequence will be positive only at one color, whereas those with two sequences will be positive for two colors. These populations can then be separated by gating the fluorescence measurements to recover single- or double-positive droplets. To demonstrate the ability to multiplex the reaction, we synthesize primers and Cy5 TaqMan probes targeting a genomic region near the 5' end of the lambda genome, and others targeting regions at increasing distances away from the 5' end. The primer and probe sequences are listed in Table 5.1, and a graphical representation of the probe locations with the Cy5 TaqMan probe in red and the FAM TaqMan probes in green is provided in Figure 5.3A. Lambda DNA is combined with the PCR reagents and the sample is emulsified using the microfluidic device. After thermal cycling, the droplets are imaged using fluorescence microscopy (Figure 5.3B) and analyzed to measure their intensity on the Cy5 and FAM channels (Figure 5.3C). The droplets are characterized as positive for both targets (Cy5+FAM+), positive for one target (FAM+Cy5-, FAM-Cy5+), or negative for both (FAM-Cy5-). Each multiplexed PCR is performed in triplicate, containing 5000-8000 droplets.

We observe less multiplexing when probe pairs are far apart, indicating that the probability that two target sequences exist within a given genome decreases for sequences that are more

separated (Figure 5.3D, blue curve); this implies that the genomes might be partially fragmented. To investigate this further, we perform a negative control in which we digest the genome with a restriction endonuclease cleaving at 10,086 bp, which is between the first and second FAM probes. If fragmentation is the source of lowered multiplexing, then the fraction of double-positives should fall precipitously beyond the cleavage point; indeed, this is what we observe, as shown by the red curve in Figure 5.3D. As an additional negative control, we digest the lambda genome using a non-specific endonuclease (Fragmentase) producing ~500 bp products, and observe that double-positives are rare for all probe pairs (green curve). This demonstrates that PAVS can be used to characterize the length distributions of viral genomes in a solution and, more generally, the presence of multiple genetic loci in a target virus, which should be useful for studying correlations between genetic loci in single viruses that are on the same linear molecule or on entirely different molecules as in segmented genomes. PAVS can also be used to characterize the integrity of viral genomes.

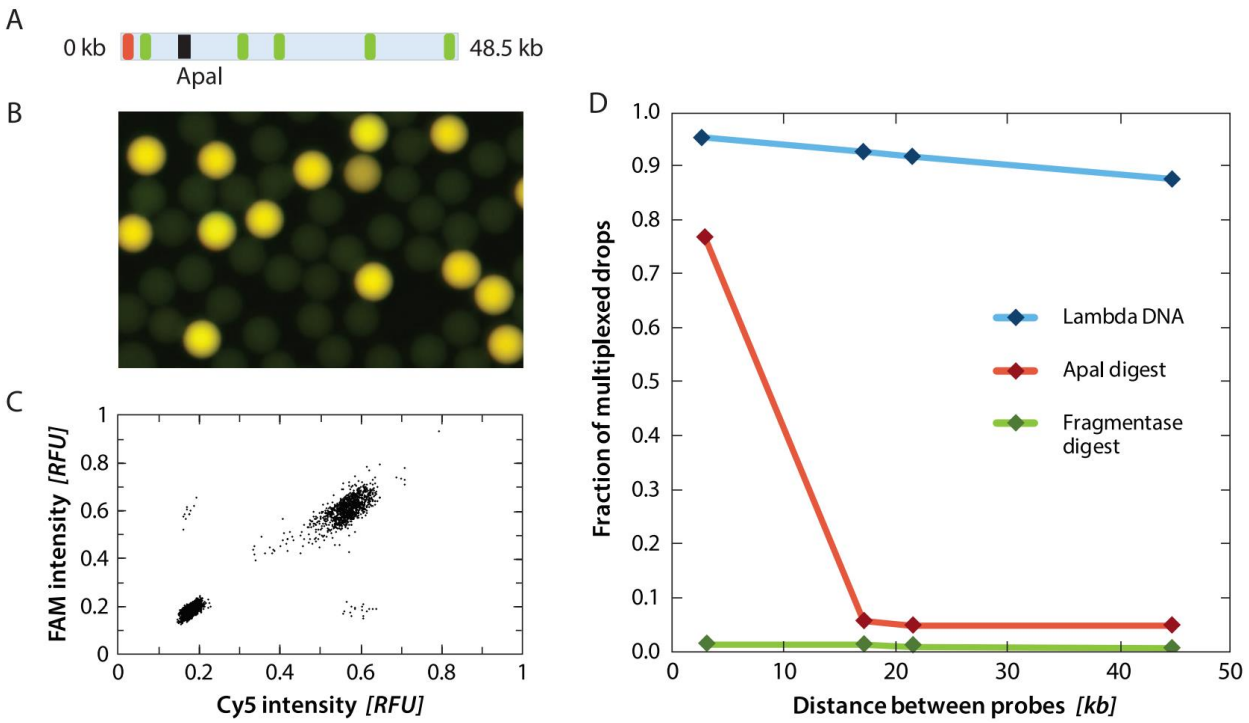


Figure 5.3 Multiplexed digital PCR detects full-length viral genomes. A) Location of TaqMan PCR Cy5 probe in the lambda genome is shown in red, FAM probes are shown in green, and the location of the ApaI restriction site is shown in black. B) Representative image of multiplexed PCR emulsion on Lambda DNA. C) Representative scatterplot of Cy5 and FAM intensities for Lambda DNA. D) Fraction of multiplexed droplets for Lambda DNA undigested (blue), ApaI digested (red), and Fragmentase digested (green).

PAVS allows target virus to be sorted out of a mixed population

The PAVS workflow consists of two steps, a first in which target viruses in a sample are detected using single virus PCR in double emulsion droplets, and a second in which the droplets are sorted to recover the target viruses. To demonstrate this, we construct a mixed sample of two bacteriophages, T4 and ϕX174, at a ratio of 1:999 respectively. This 0.1% T4 spike-in is

encapsulated at limiting dilution in double emulsions with PCR primers specific for T4 phage, thermally cycled, and stained with SYBR Green. If a particular droplet contains T4, the nucleic acids targeted by the PCR primers are amplified and the SYBR stain produces a fluorescent signal that fills the droplet. The fluorescence signal is detected using FACS and the positive droplets sorted into an 1.5 mL microcentrifuge tube.

To validate that the PAVS workflow enriches for T4 over ΦX174, we quantify the virus concentrations in the sorted and unsorted pools using qPCR. The sorted droplets are ruptured and the viral genomes they contain amplified by ddMDA. Equal concentrations of T4 and ΦX174 DNA from the unsorted and sorted emulsions are subjected to qPCR (Figure 5.4). The primers used to detect T4 target a different locus than the ones for PAVS sorting (Table 5.1). The qPCR curve for T4 shifts to lower cycles post-sorting, demonstrating that T4 has been enriched. By contrast, the curve shifts to higher cycle numbers for ΦX174, indicating that this virus has been de-enriched by sorting, as expected. To quantify the degree of enrichment, we compute an enrichment factor *e* defined as,

$$e = \frac{(n+1)\left(\frac{1}{2^{\Delta C_t^{\mathrm{T4}}}}\right)}{\left(\frac{1}{2^{\Delta C_t^{\mathrm{T4}}}}\right) + n\left(\frac{1}{2^{\Delta C_t^{\phi\mathrm{X174}}}}\right)},$$

where n is the ratio of the viral species with respect to one another and ΔC_t^{T4} and $\Delta C_t^{\phi\mathrm{X174}}$ are the differences of cross-threshold values for T4 and ϕX174, respectively. For this experiment, n = 999, ΔC_t^{T4} is 2.16, and $\Delta C_t^{\phi\mathrm{X174}}$ is 5.45, yielding e = 9.69, indicating that the final sample is enriched by about tenfold for T4 from an initial concentration of 0.1%. Larger enrichments can be

achieved by further diluting the sample prior to partitioning into droplets, which reduces the rate of co-encapsulation of the two species and false-positive recovery of off-target cells.

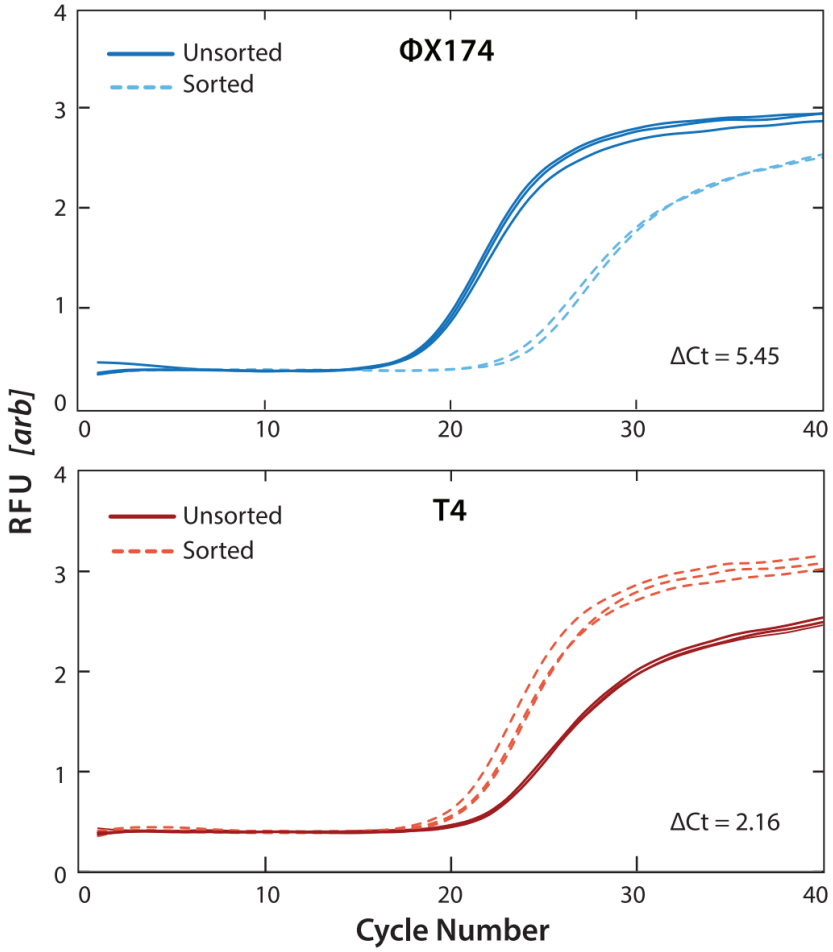


Figure 5.4 qPCR detection of bacteriophages ϕX174 and T4 before and after FACS sorting. The shifts in the curves reflect the 2-fold change of the DNA quantity according to the specific primers being tested. Samples tested in triplicate.

PAVS recovery of single virions from a mixed sample

Common FACS instruments can pool all positive droplets into one well or dispense controlled numbers into different wells, including down to single droplets. This is commonly used

to isolate cells for single cell analysis (Figure 5.5A). When combined with PAVS, this allows a heterogeneous mixture of viruses to be sorted, to isolate specific virions in the sample, which can then be subjected to additional analyses, such as qPCR. To illustrate this, we sort a sample of lambda virus with PAVS and dispense the positive droplets into wells in controlled numbers (Table 5.1, Lambda FWD 2, Lambda REV 2, and Lambda probe 2). We load 1, 10, or 50 positive droplets into each well and analyze the recovered material with qPCR for primers targeting a different portion of the lambda genome than was amplified in the PAVS detection (Table 5.1). The C_t values decrease as the number of viruses dispensed increases, indicating that the viruses are present at higher numbers (Figure 5.5B). When fewer than 50 viruses are sorted, it is difficult to reliably detect them in the sorted wells; wells with 10 viruses show amplification at C_t values of 33, while single viruses do not amplify above the negative controls.

To confirm that the sorting is specific, we generate qPCR curves for wells containing 50 positive droplets and wells containing 50 unsorted droplets. The qPCR curve shifts left by an average of 4.24 C_t values, demonstrating that the Lambda virus is more abundant in the sorted sample (Figure 5.5C). While our results show that single viruses provide too little DNA for detection with standard qPCR, other post-sorting amplification methods may be implemented to improve sensitivity, such as nested PCR[264] or non-specific ddMDA followed by qPCR[263].

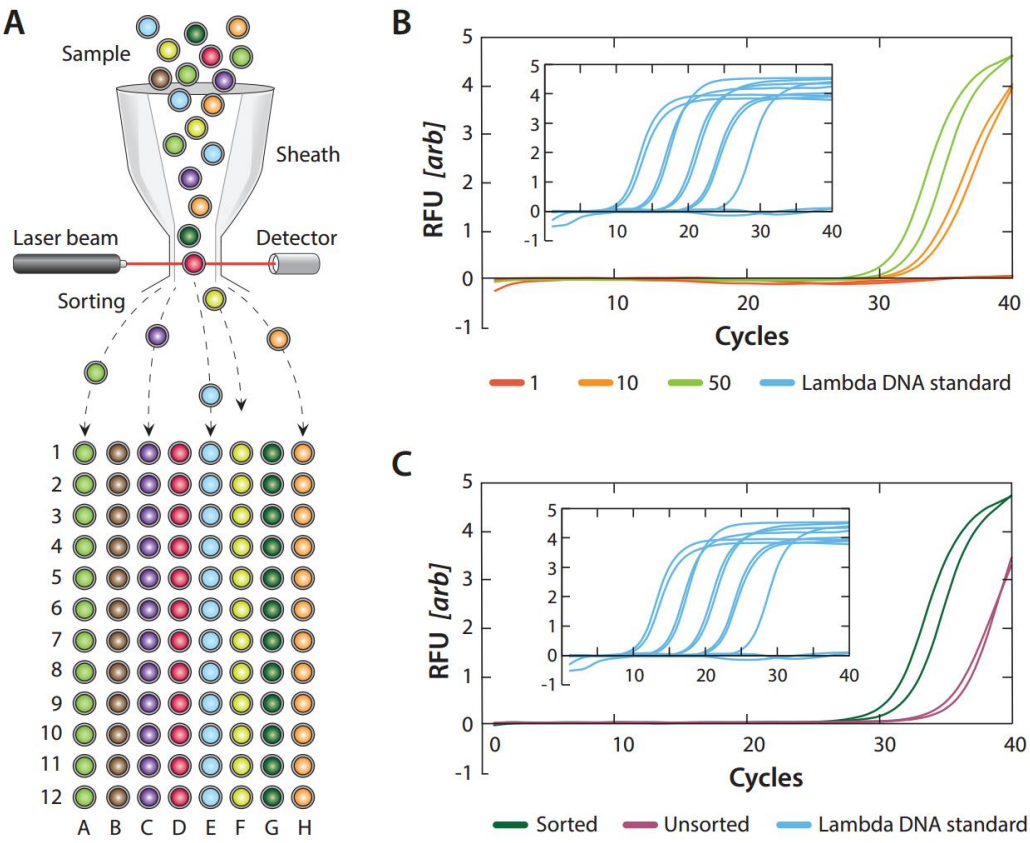


Figure 5.5 PAVS recovery of single virions from a heterogeneous sample. A) Workflow schematic for targeted virus sorting into well plates using PAVS. B) qPCR detection of Lambda for 1, 10, or 50 viruses sorted into a well. Positive control of Lambda DNA shown in inset. C) qPCR curves for 50 sorted or unsorted Lambda viruses per well. Positive control of Lambda DNA shown in inset. Samples tested in duplicate.

5.5 Conclusions

PCR-activated virus sorting enables specific detection and sorting of viruses in a mixed sample, analogous to what is possible with FACS for cells. Just as when studying microbes, specific and high throughput virus sorting should be valuable whenever large populations of virus must be analyzed to recover a target species, provided ~100 bp can be defined for PAVS detection. It can also be used to detect viruses residing within host microbes, including bacteria and

eukaryotic cells. This should be valuable for characterizing virus-host relationships in microbial ecologies or the human microbiome, for example, to determine which viruses infect which hosts – something that is presently extremely challenging due to the inability to culture most viruses and host microbes. PAVS opens new avenues in the study of virus-mediated diseases, such as HIV and EBV, where it enables detection of viral infection within host cells. The recovered cell lysates can then be sequenced to characterize, for example, insertion sites of the virus, epigenetic correlations with virus infection, or modulation of host cell transcriptome, all of which should be valuable for studying the basic biology of the virus and to better understand how it survives, replicates, and evades the host immune response.

5.6 Supplemental Information

Primers for PCR	**Sequence**
T4 probe	5'-/56-FAM/TCGCATTCT/ZEN/TCCTCTGATGGAGCA /3IABkFQ/ -3'
T4 FWD	5’-CCACAACTAACCGAGGAAGTAA-3’
T4 REV	5’-TGCGATATGCTATGGGTCTTG-3’
Lambda probe 1	5’- /5Cy5/TCTGCCCGTGTCGGTTATTCCAAA/3IAbRQSp/-3’
Lambda FWD 1	5’- GACTACATCCGTGAGGTGAATG-3’
Lambda REV 1	5’- TTGCGCTGCTTATGCTCTAT-3’
Lambda probe 2	5’- /56-FAM/TGGGTGTCTCGTATGAGCAGCTTT/3IABkFQ/-3’
Lambda FWD 2	5’- TAACGGCTACTCCGTGTTTG-3’
Lambda REV 2	5’- AGCTCATCTGGGCGTAATTC-3’
Lambda probe 3	5’- /56-FAM/TGAGCAGAC/ZEN/CAAAGACGGCAAACA/3IABkFQ/-3’
Lambda FWD 3	5’- TGGAAGGATGCCAGTGATAAG-3’
Lambda REV 3	5’- TCCATGCTGAGGCCAATA C-3’
Lambda probe 4	5’- /56-FAM/CTCGTTGCT/ZEN/GGAAGCCTGGAAGA/3IABkFQ/-3’
Lambda FWD 4	5’- CAGGTAGCCAGTGAGCATATT-3’
Lambda REV 4	5’- GTTCAGCAACACCCGATACT-3’
Lambda probe 5	5’- /56-FAM/ATACTGAGC/ZEN/ACATCAGCAGGACGC/3IABkFQ/-3’
Lambda FWD 5	5’- GCCCTTCTTCAGGGCTTAAT-3’
Lambda REV 5	5’- CTCTGGCGGTGTTGACATAA-3’
Lambda probe 6	5’- /56-FAM/TATCCGTCA/ZEN/GGCAATCGACCGTTG/3IABkFQ/-3’
Lambda FWD 6	5’- GTGGCATTGCAGCAGATTAAG-3’
Lambda REV 6	5’- GGCAGTGAAGCCCAGATATT-3’
Primers for qPCR	**Sequence**
T4 FWD	5’-ACCCGGACCAAAATCTCGAC-3’
T4 REV	5’-GCGCAGTAGTCCGTGAATTG-3’
ΦX174 FWD	5’-TTCTGTGCCGCGTTTCTTTG-3’
ΦX174 REV	5’-AAACAGGGTCGCCAGCAATA-3’
Lambda FWD	5’-CGTGATGGAGCAGATGAAGAT-3’
Lambda REV	5’-GTATCCAGCTCACTCTCAATGG-3’

Table 5.1 Primers and TaqMan probe sequences

Chapter 6: PCR-activated cell sorting for characterizing virus-host relationships

The following section is reprinted from the submitted manuscript, "PCR-activated cell sorting for characterizing virus-host relationships" by Shaun Lim, Shea T. Lance, Kenneth M. Stedman, and Adam R. Abate. Shaun Lim, Shea Lance, Kenneth Stedman, and Adam Abate designed the research. Shaun Lim and Shea Lance performed experiments. Shaun Lim, Shea Lance, Kenneth Stedman, and Adam Abate wrote the manuscript.

In this chapter, we describe the application of PCR-Activated Cell Sorting (PACS) for detection and enrichment of specific viruses out of a heterogeneous sample in order to identify the viral host species. This application is similar to the PAVS technique presented in Chapter 5, except this treatment uses viral detection and enrichment to identify the viral hosts. Characterizing viral hosts using conventional methods is challenging because most microbes cannot be cultured and most viruses cannot be cultured. PACS does not require the target cells or virus to be cultivable, and it is ultrahigh-throughput to allow detection of rare viral species. In this chapter, we demonstrate the enrichment of one viral species and the identification of the viral host using qPCR.

6.1 Abstract

Characterizing virus-host relationships is critical for understanding the impact of a virus on ecosystems, but is challenging with existing techniques, particularly for uncultivable viruses and hosts. We present a general, cultivation-free approach for characterizing virus-host relationships. Using picoliter droplet-based microfluidic PCR-activated cell sorting, we interrogate millions of individual cells in a heterogeneous population for the presence of virus nucleic acids. Hosts harboring virus nucleic acid are recovered and identified.

6.2 Introduction

Viruses substantially impact the health and dynamics of all communities, from causing disease in individuals to influencing global biogeochemical cycles[246,247]. Cyanophages, for example, play an important role in oceanic carbon fixation, since a core photosynthetic protein is of phage origin[265,266]. Moreover, it has become clear that there are an astronomical number of viruses in the biosphere[247], but it is unclear what hosts the vast majority of these viruses infect. Studying the impact of environmental viruses necessitates methods for characterizing virus-host interactions. However, the gold-standard, the plaque assay, requires that the host cell be cultivable and that virus infection generate a unambiguous phenotype; this is not the case for most virus-host pairs, particularly for microbes and their viruses [267,268].

Next generation sequencing (NGS) is a powerful, culture-independent method for studying virus-host interactions in diverse population[269–272]. In this approach, metagenomic sequences are interrogated for molecular traces left by the viruses co-occurring with clear host genes such as small subunit ribosomal DNA sequences. These virus traces are often resistance-conferring sequences, such as CRISPR-associated sequences[273,274]. Identifying virus-host relationships using NGS is only possible, however, if such traces are identifiable. Moreover, even when they occur, these sequences are rare, making their identification even more challenging. Single cell genomics of uncultivated cells is another powerful method for identifying new virus-host pairs[271,272]. In one study, 127 single amplified genomes (SAGs) from a known clade of uncultivated organisms were screened for viruses and 69 new virus clades were found[271]. In another study, bioinformatic analysis of 58 SAGs allowed the identification of 30 new virus genomes[272]. The isolation of single cells for SAG analysis with FACS is non-specific, however: Cells of interest are identified only after sorting using PCR or other sequence analysis. Since the throughput of the method is low, it is challenging to identify the hosts of rare viruses.

Another strategy for identifying virus-host relationships is to sort cells based on viral infection prior to sequencing. For example, PhageFISH (Fluorescence In Situ Hybridization) allows cells to be separated based on virus infection by marking infected cells with fluorescent probes[275]. PhageFISH is high throughput since millions of cells can be stained in parallel and sorted at >10,000 per second with fluorescence-activated cell sorting (FACS). However, probe hybridization is performed in the complex milieu of a fixed cell, necessitating substantial assay optimization that is extremely challenging with uncultivable organisms. Viral genome tagging, on the other hand, labels the viral genomes so that host cells become fluorescent upon infection, allowing them to be recovered via sorting[276]. Labeling viral genomes, however, is nontrivial for uncultivable viruses.

A powerful, cultivation-free method for identifying specific virus hosts is single cell PCR in individual microfluidic chambers[277]. In this approach, TaqMan PCR interrogates individual microbes for virus sequences such that, if the sequences are present, a fluorescent signal is produced in the chamber. This allows identification and recovery of infected cells by interrogating the positive chambers. Because this method uses PCR, which yields exponential amplification when virus sequences are present, the signal difference between infected and uninfected cells is large, permitting unambiguous identification. Moreover, harsh thermocycling facilitates cell lysis and enhances access of hybridization primers to their targets, reducing false-negatives. The principal limitation of this method is that it lacks scalability, enabling interrogation of just thousands of cells; this limits its utility for studying most environments, comprising billions of microbes per milliliter[278]. To allow comprehensive characterization of virus-host relationships, an optimal method would be applicable to uncultivable species, capable of analyzing millions of cells, and enable the specific recovery of all cells infected with the target virus.

We present a general method for identifying virus-host relationships that is applicable to uncultivable organisms; the method combines the scalability of FACS with the generality and precision of single cell PCR. Using PCR-activated cell sorting (PACS) we sort cells based on infection by specific viruses[279,280]. This is accomplished by isolating individual cells from a heterogeneous population in water-in-oil droplets and performing PCR in each droplet with primers specific to the target virus genome (Figure 6.1). If the virus is present either in the host or bound to it, the PCR produces a fluorescent signal, allowing the host cell and virus genomes to be recovered by sorting the encapsulating droplet. The sorted genomes are then analyzed to identify the host species.

PACS has a number of advantages over other methods for characterizing virus-host relationships: It does not require the target cells or virus to be cultivable, making it general. It is ultrahigh-throughput, allowing millions of individual cells to be sorted based on virus infection, permitting detection of cells infected with rare viruses. The TaqMan assay can be multiplexed using probes of different color allowing, for example, recovery of cells in a sample containing multiple, specific sequences, such as due to co-infection by distinct viruses. This method is particularly applicable to the identification of hosts for viruses whose sequences have been identified in metagenomic studies.

6.3 Materials and methods

Preparation of bacteriophages, plaque assays and cells.

Bacteriophage T4 (T4), bacteriophage ΦX174 (ΦX174), and *E.coli* hosts were obtained from Carolina Biological Supply. T4 was propagated by infection of *E.coli* B (ATCC 11303)[281] and ΦX174 by infection of *E.coli* C (ATCC 13706). Bacteriophage lambda (lambda cI857^{ts}) was obtained from the lambda lysogen *E.coli* strain KL470 graciously provided by R. Raghavan.

Bacteriophage lambda was propagated by infection of *E.coli* C600 (Carolina) and plaques formed as previously described[282]. A lambda lysogen of lambda cI857^{ts} in *E.coli* C was prepared[282] and purified by two rounds of single colony isolation on LB agar. Titers of T4, ΦX174 and lambda were determined with plaque assays as previously described[283] (Karam et al, 1994; Arber et al. 1983). Bacteriophage-containing lysates were separated from cellular debris by 5 minutes of centrifugation at 3000 x g, filtered through 0.2 µm filters (Sartorius, Minisart) and preserved with a single drop of chloroform in the preparation.

Microfabrication of devices

The microfluidic chips are fabricated using standard photolithography techniques in poly(dimethylsiloxane) (PDMS)[201]. To produce a device master from which the PDMS replicates are molded, SU-8 photoresist (MicroChem) is spun onto a 3" silicon wafer at a thickness of 25 µm, and exposed to UV light from a UV photodiode (ThorLabs) through a UV-absorbent Mylar mask containing an inverse-image of the microfluidic chip (Fineline Imaging). The wafer is baked at 95ºC on a hotplate for 1 minute, and developed in propylene glycol monomethyl ether acetate (PGMEA) to remove uncrosslinked resist, followed by post-baking in accordance with the manufacturer's instructions. PDMS polymer and crosslinker is combined at a ratio of 11:1, poured over the master, degassed to remove trapped air bubbles, and baked at 75ºC for 4 hours to crosslink the device. The device is peeled from the master and holes are punched using a 0.75 mm biopsy coring needle. The punched device is washed with isopropyl alcohol and patted with scotch tape to remove debris prior to plasma bonding to a glass slide. To render the channels hydrophobic for water-in-oil emulsification, Aquapel^{TM} is flushed into the channels, after which the device is baked in an oven for 20 min at 65ºC.

Encapsulation of sample in microfluidic droplets

Prior to encapsulation, bacterial suspensions are washed three times by centrifugation at 3000g (Eppendorf microcentrifuge) and re-suspension of the pellets in distilled water. Viral suspensions are encapsulated without washing. The suspensions are mixed with primers, TaqMan probes, and PCR mix (2X ddPCR MasterMix, Bio-rad). The primers and TaqMan probes are used at 1 µM and 250 nM, respectively. The mix is loaded into a 1 ml syringe atop 200 µl HFE-7500 fluorinated oil (3M), connected to a PDMS flow-focus droplet generator through a 21 gauge needle and polyethylene tubing. Droplet generation oil for probes (Bio-Rad) is introduced into the carrier-phase inlet of the microfluidics device through another syringe and tube; the oil comes with a proprietary surfactant included to stabilize the generated droplets during the heating and cooling steps involved in PCR.. Using syringe pumps, the aqueous phase is injected at 200 μlhr^{-1} and the oil at 400 μlhr^{-1} (New Era), generating 25 µm diameter droplets at ~3.6kHz in a droplet maker with nozzle 20 µm wide and 25 µm tall. The emulsion is collected into 200 µl PCR tubes and thermocycled on a T100 thermocycler (Bio-Rad), using the following conditions: 10 min. at 95^{o}C, 35 cycles of 10 s. at 95^{o}C, 15 s. at 55^{o}C and 30 s. at 70^{o}C. To verify specificity of the PCR, the emulsions are chemically ruptured with chloroform and DI water and the aqueous fractions electrophoresed on a 2% agarose gel to confirm correct amplicon length.

Detection and sorting of droplets

After thermocycling the emulsions, the droplets must be sorted based on fluorescence. This is accomplished by loading the thermocycled emulsions into a syringe with 200 µl HFE-7500, maintaining the syringe vertically so that the needle faces up, and allowing the emulsion to cream

for ~10 min; this ensures the droplets are at the top of the emulsion before the syringe pump is started so that they flow into the device at a controlled flow rate and are closely packed. The droplets are injected into the detection and sorting device (Fig. 1)[143,279,280,284] at a flow rate of 50 μlhr^{-1}, with spacer oil flow rate 1000 μlhr^{-1}. The flow rate for the second oil spacer at the sorting junction is set to 100 μlhr^{-1}. All droplet spacing is performed with pure HFE-7500. All electrodes on the device, including the sorting electrode and the moat that shields the droplets from stray field, are filled with 2M NaCl solution[285,286]. A 100 mW, 532 nm laser is focused upstream of the sorting junction to excite droplet fluorescence. Photomultiplier tubes (PMTs) focused on the same spot measure emitted fluorescent light and output a voltage proportional to the light intensity to a computer outfitted with an FPGA card (National Instruments) programmed in LabVIEW. The card detects droplets as peaks in fluorescence over time and, when a droplet is to be sorted, outputs a 40 kHz, signal amplified to 1000 V (Trek) applied to the salt-water electrodes on the microfluidic chip. Custom LabVIEW software allows adjustment of PMT gain, droplet fluorescence intensity thresholds for sorting, and electrode AC voltage pulse frequency and magnitude.

Quantitative PCR analysis of sorted droplets

Genomic material from the sorted drops are recovered by rupturing the drops via addition of 100 µl chloroform together with 50 µl DI water and vortexing for 10 min[279]. To measure the degree of enrichment via PACS, we probe for specific genomic regions in both E. coli strains (K-12 and C) before and after sorting. For E. coli C, we use primers specific to PRP, a gene that has been shown to be in that strain but not in E. coli K-12. genome. For E. coli K-12, the strain that was used in this paper has an integrated ybgF-mCherry cassette, and primers specific to mCherry were used. To test the specificity of the primers, the primers for *E. coli* K-12 and C are tested

against each other's' target templates; we observe no background amplification using qPCR or by interrogating the amplified products with gel electrophoresis. The primers are tested for linearity by constructing a serial dilution curve for each primer set over a hundred thousand-fold degree of dilution for the target template. qPCR measurements of the genomic material yielded cross-threshold numbers that correspond to the number of PCR cycles needed for fluorescence levels to move past a predetermined fluorescence threshold level. We obtain 2 cross-threshold numbers for each of the pre- and post-sorted samples, and via calculations as detailed in this paper, obtain an enrichment factor that is indicative of the degree of enrichment of the pre-sorted sample. The amplification reagent for all the qPCR measurements is Maxima SYBR Green Master Mix (Thermo Scientific).

6.4 Results

PACS workflow for identifying virus-host relationships

To enable the ultrahigh-throughput sorting of microbes based on the presence of viral nucleic acids, we use PACS, a droplet-based microfluidic technology[279,280]. In PACS, picoliter-volume aqueous droplets are used as reactors in which to perform TaqMan PCR on single cells. Using flow-focus emulsification[54,66,287], individual particles or cells from a heterogeneous sample are isolated in ~10^7 droplets with PCR reagents and probes targeting specific viral genes (Figure 6.1). At a loading rate of ~0.1 cells per droplet, we interrogate 10^6 cells. After all cells are encapsulated, the emulsion is thermocycled, performing 10^6 parallel single cell PCR reactions, as illustrated in Figure 6.1B. During thermocycling, the cells lyse and their nucleic acids are subjected to TaqMan amplification. If a particular droplet contains the nucleic acids of a virus targeted by the probes, the nucleic acids are amplified, generating a fluorescent signal that fills the encapsulating droplet. This marks the droplet as containing a cell infected with the target virus,

allowing the host cell nucleic acids to be recovered by sorting the droplet, as illustrated in Figure 6.1.

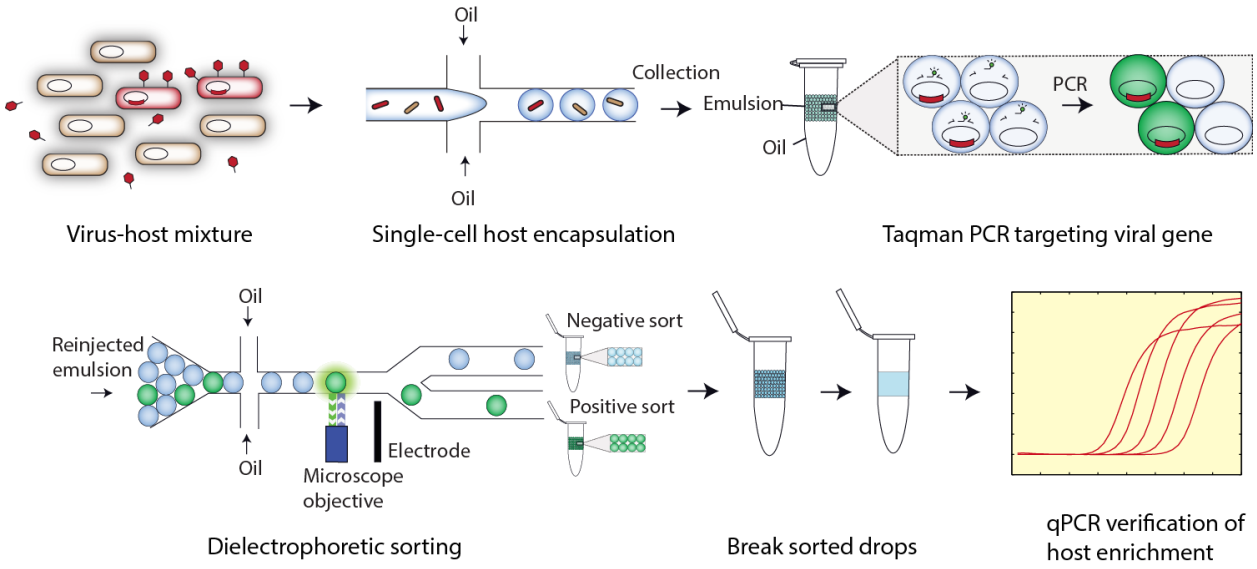


Figure 6.1. Microfluidic workflow for PACS-based viral detection and host sorting. Virus-infected hosts are first encapsulated with PCR reagent, primers and probe in picoliter-volume droplets, then thermocycled to yield fluorescent drops. These drops contain targeted genomes of viruses and their hosts, which are then sorted to yield a purified population of DNA. This material can be used for downstream sequencing analysis in this case qPCR, but any sequence analysis is possible.

Specific detection and quantification of viral genomes from bacteriophage T4 and ΦX174

Detecting cells infected with a specific virus depends on the ability to reliably and specifically amplify virus nucleic acids in infected cells. To investigate the robustness of this step in the PACS process, we perform TaqMan PCR in microfluidic droplets on two distinct virus species, bacteriophage T4 (T4) and bacteriophage ΦX174 (ΦX174). Pure preparations of the viruses are combined with PCR and TaqMan reagents immediately prior to microfluidic

emulsification and thermocycling. After thermocycling, we observe clearly fluorescent droplets in a population of non-fluorescent droplets, as shown in Figure 6.2A, indicating successful amplification of the viruses. In negative controls, we perform the same experiment but swap the TaqMan probes by including T4 probes in ΦX174 preparations and ΦX174 probes in T4 preparations. Neither of these controls yields detectable fluorescent droplets, demonstrating that the probes are specific to their target virus.

An important factor when using PACS to enrich cells out of a heterogeneous sample is the rate of false negatives since this limits the number of positive events detected. To characterize the sensitivity of the method, we scan the emulsions created in the previous experiment using flow dropometry[279], recording fluorescence values for ~30,000 individual droplets (Figures 6.5, 6.6). Using the known droplet volumes and assuming that virus encapsulation is governed by Poisson statistics[76], we estimate virus concentrations in the starting samples and compare them to estimates from plaque assays, Figure 6.2B. For bacteriophage T4, plaque assays on samples used for PACS yields $3.0 * 10^9$ pfu/ml compared to the $3.1 * 10^9$ particles/ml for PACS. For ΦX174, the plaque assay yields $3.0 * 10^9$ pfu/ml versus $3.5 * 10^9$ particles/ml for PACS. The estimates using both methods are in excellent agreement indicating that the PCR conditions are sufficient to efficiently lyse the viral particles. The slightly higher values estimated by PACS may reflect that some virus genomes are incorrectly packaged, mutant, or are in non-infectious particles. In addition to validating droplet PCR for PACS, this demonstrates that droplet digital PCR is an alternative approach for quantitating virus genomes in a sample.

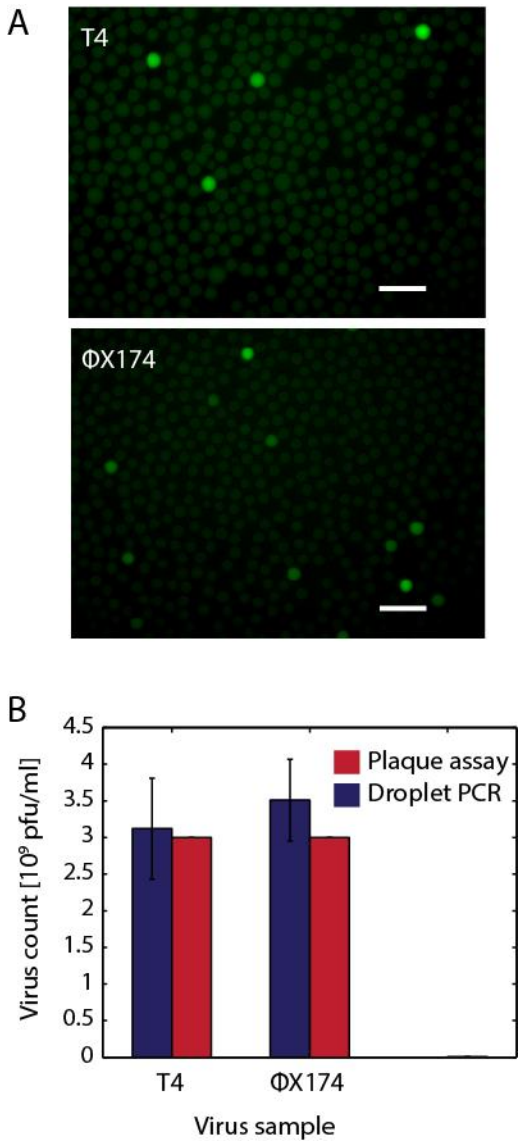


Figure 6.2 Specific detection and quantification of viral genomes from bacteriophage T4 and ΦX174. A) Digital detection of viral particles after droplet PCR. Bacteriophages T4 and ΦX174 virions are partitioned into droplets for TaqMan PCR detection Scale bars are 100µm. B) Plaque assay results closely mimic digital viral particle quantitation, suggesting that viral genomes are accurately measured with this new method. Error bars represent the standard deviation of 3 technical replicates.

Sorting E. coli infected with lambda bacteriophage

PACS enables the detection of cells infected with specific viruses, including lysogens, and recovery of host cell genomes. To illustrate this, we construct a test system comprising two *E. coli* strains: C, a strain that lambda virus can infect, and K12, a strain that it cannot. To validate the TaqMan probes, we analyze a sample of pure virus in suspension and observe digital signals among the droplets, indicating single-virion detection (Figure 6.3A). We then spike in *E. coli* C. infected with lambda into uninfected *E. coli* K-12 in a ratio of 1:9, thereafter washing the mixed suspension to remove free virus. We subject this sample to droplet PCR and again observe a digital signal corresponding to a small subpopulation of positive droplets which, presumably, contain *E. coli* cells infected with lambda virus, as shown in Figure 6.3B. To verify that the digital fluorescence corresponds to droplets containing lambda infected cells, we sort the emulsion to recover the positive droplets, which we accomplish using dielectrophoretic droplet sorting, Fig. 1[143,279,280,284,285]. The thermocycled emulsion is injected into the sorting device, which flows the droplets spaced by oil individually through the focused excitation laser. As a droplet passes through the laser, its fluorescence is excited and the resulting emitted light is measured with a photomultiplier tube (PMT), which outputs a voltage proportional to the fluorescence intensity analyzed by the computer and FPGA card. The droplets appear as peaks in voltage as a function of time, in which the amplitude of the peak is proportional to the droplet intensity, as shown in Figure 6.3C. When a positive droplet passes through the laser, an abnormally tall peak is observed, as seen at $t = 0.0185$ seconds (Figure 6.3C). Upon detection of a positive droplet, the computer outputs an alternating voltage amplified to ~1000 V applied to on-chip electrodes generating an attractive force that pulls the positive droplet into the collection channel[143,279,280]. When a negative droplet passes through the laser the peak amplitude does not fall above the user-defined threshold;

the electrode remains un-energized and the droplet flows passively into the waste channel. In this way, the droplet sorter separates the positive from the negative droplets.

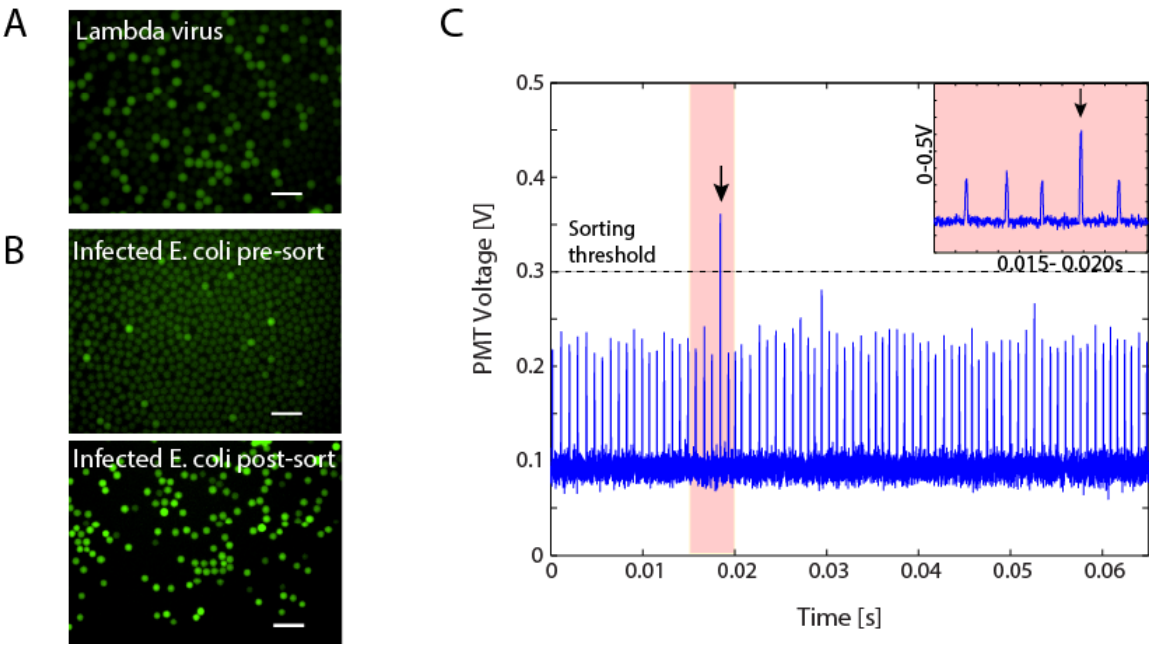


Figure 6.3 Sorting E. coli infected with lambda bacteriophage. A) Digital detection of lambda particles and B) pre- and post- sorted drops containing lambda and its E. coli host. Digital detection of lambda using the probes and primers from Figure 3A. C) Time trace of fluorescent droplet detection. Droplets are run through a dielectrophoretic microfluidic sorter, and droplets above a set threshold are sorted and collected each peak represents a single droplet. Inset figure shows a magnified view of the lone sorted event in the time period 0.015-0.025 seconds. All scale bars are 100µm.

Enrichment of phage-infected E. coli genomes

Sorting the cells based on presence of lambda DNA should yield primarily *E. coli* C cells, since only this strain can be infected by lambda virus. To quantify the enrichment afforded by PACS,

we analyze the fraction of *E. coli* C DNA in the sorted population using qPCR. We perform a second sort in which we lower the concentration of *E. coli* C to 0.1%, while increasing *E. coli* K12 to 99.9%. Using PACS, we recover all TaqMan positive droplets and the nucleic acids from their cells. To measure the fraction of nucleic acids corresponding to *E. coli* C and *E. coli* K12, we perform qPCR on the nucleic acids pre- and post-sorting using primers specific to the two strains; the results for both the 10% and 0.1% mixtures for both primer sets are provided in Figure 6.4. The signal from *E. coli* K12 is strongly *de-enriched* in the PACS-sorted material, corresponding to curve shifts by ~10 C_t values (Figure 4, upper panels). By contrast, the curves for *E. coli* C shift in the opposite direction, demonstrating that PACS for lambda virus nucleic acids *enriched* for these cells. To obtain a quantitative metric of enrichment, we define the enrichment factor as the ratio of host microbe purity in the pre-sorted to the post-sorted samples,

$$e = \frac{C_{post}(C_{pre}+K_{pre})}{C_{pre}(C_{post}+K_{post})},$$

where C_{pre} and C_{post} are the number of *E. coli* C cells present in the pre- and post-sorted samples, while K_{pre} and K_{post} are the number of E.coli K-12 present in the pre- and post-sorted samples, respectively. The relationships between C_{pre} and C_{post}, together with K_{pre} and K_{post}, can be determined via qPCR, which measures the log-2 fold change of gene copy numbers specific to either microbe,

$$C_{post} = (2^{\Delta C_{tc}})C_{pre} \text{ and } K_{post} = (2^{\Delta C_{tk}})K_{pre},$$

where ΔC_{tc} and ΔC_{tk} are the differences between the qPCR cross-threshold values for *E. coli* C and K-12, respectively. This is an estimation since the actual rate of DNA amplification may be less than 2 per cycle. We know C_{pre} and K_{pre} because we begin with controlled numbers of each species before sorting, enabling us to define the ratio of the two species with respect to one another,

$$n = \frac{K_{pre}}{C_{pre}},$$

and to simplify the enrichment factor to,

$$e = \frac{(n+1)\left(\frac{1}{2^{\Delta C_{tc}}}\right)}{\left(\frac{1}{2^{\Delta C_{tc}}}\right) + n\left(\frac{1}{2^{\Delta C_{tk}}}\right)}.$$

For our initial sort, $n = 9$ (10% *E. coli* C to K12 ratio), we obtain 4.0 for ΔC_{tc} and 13.1 for ΔC_{tk} (Fig. 4.), yielding $e = 9.84$, indicating that the final sample is enriched to 98.4% for *E. coli* C from an initial concentration of 10%. For the 0.1% spike-in, we obtain 3.0 for ΔC_{tc} and 9.9 for ΔC_{tk} (Fig. 4.), providing $e = 106$, so that *E. coli* C is enriched by about a hundred times to 10.6% final concentration. Larger enrichment factors can be achieved by further diluting the sample prior to partitioning into droplets, which reduces the rate of co-encapsulation of the two species and false-positive recovery of off-target cells.

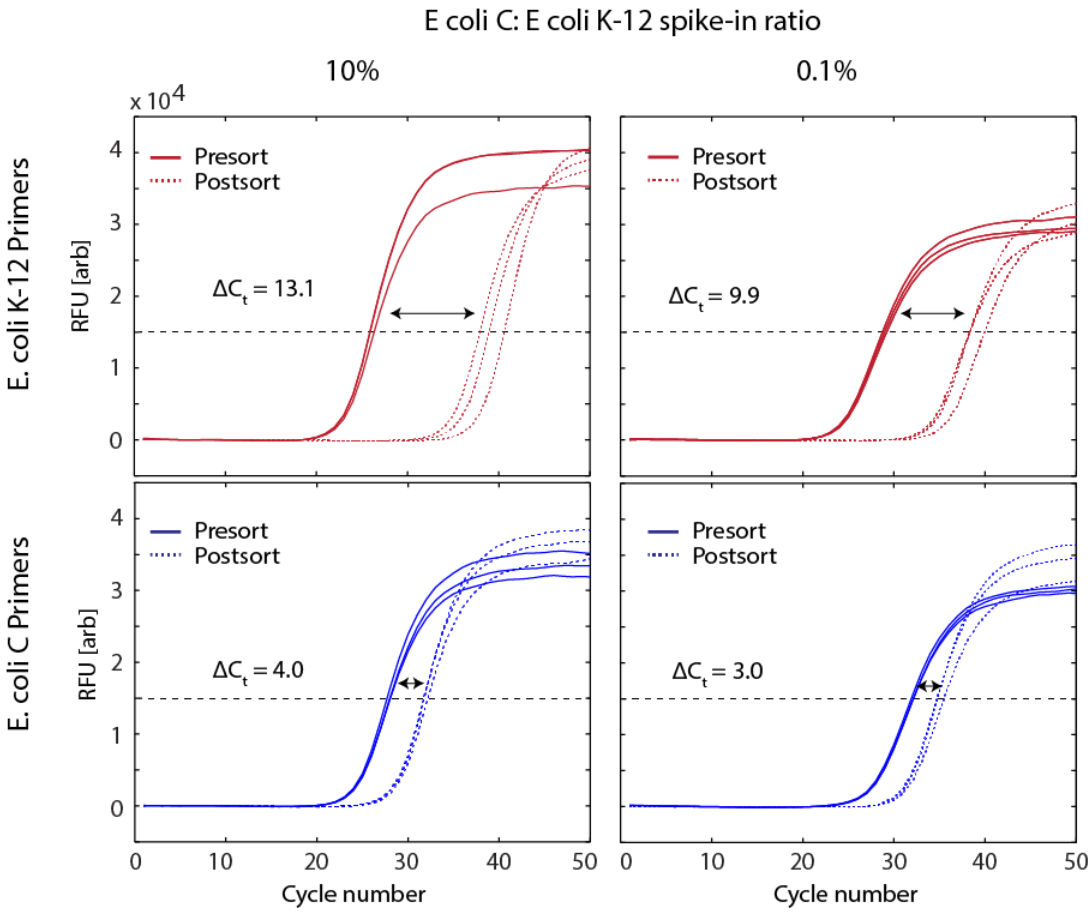


Figure 4. qPCR detection of lambda host genomes before and after droplet sorting. Each quadrant shows qPCR amplification curves for DNA extracted from drops before and after sorting. The shifts in the curves reflect the 2-fold change of the DNA quantity according to the specific primers being tested.

6.5 Discussion

When confronted with large, heterogeneous populations, cell sorting is invaluable because it allows subpopulations to be systematically isolated and studied. However, currently, specific cell sorting is only possible using affinity reagents, such as antibodies or oligomer probes that specifically bind to the cell type of interest[224,288]; such reagents are rarely available for uncultivable targets. Nonetheless, in many of these cases, sequence data are available. In these instances, PACS

is uniquely suited for specific cell sorting since it relies on PCR to differentiate between cells based on their nucleic acid content. In this application of PACS, cells are sorted based on the presence of a viral genome within or attached to the host cell (Figure 6.4). We cannot differentiate between bound virus and infected cells using techniques presented here as intact virions also provide positive PACS signals (Figure 6.2), but a stripping step could be included in the PACS workflow, if only cells with internal viruses are desired.

In its current implementation, PACS differentiates between cells based on one sequence. However, since PACS relies on TaqMan assays, probes of different color can be used to interrogate multiple sequences simultaneously[232]. This should enable multi-parametric cell sorting similar to current FACS analysis with multiple with different antibodies. Multi-parametric cell-sorting could recover, for example, all cells infected with specific variants of a virus, co-infection of different viral species, or the presence of specific host and virus sequences associated with the same cell.

A limitation of the current system is that the lysis of the cells is achieved using the detergents present in the PCR reagents and the high temperature of PCR thermocycling, which may be insufficient for lysing some cells or viruses. To broaden the applicability of the method, harsher lysis procedures could be implemented that allow the inclusion of enzymes to digest cellular material. Such procedures can be implemented using recently described agarose emulsion and multi-step droplet merger workflows, which enable the digestion of cell lysates with proteases[280]. Indeed, this is essential for reliably performing PACS on mammalian cells in sub-nanoliter droplets, since undigested mammalian cell lysates potently inhibit PCR[280].

PACS is founded on droplet-based microfluidic technologies whose intrinsic throughput is 1,000 droplets per second[287]. This enables facile and rapid sorting of millions of cells[285]. By implementing faster emulsification strategies using air-bubble triggered droplet generation[69] or

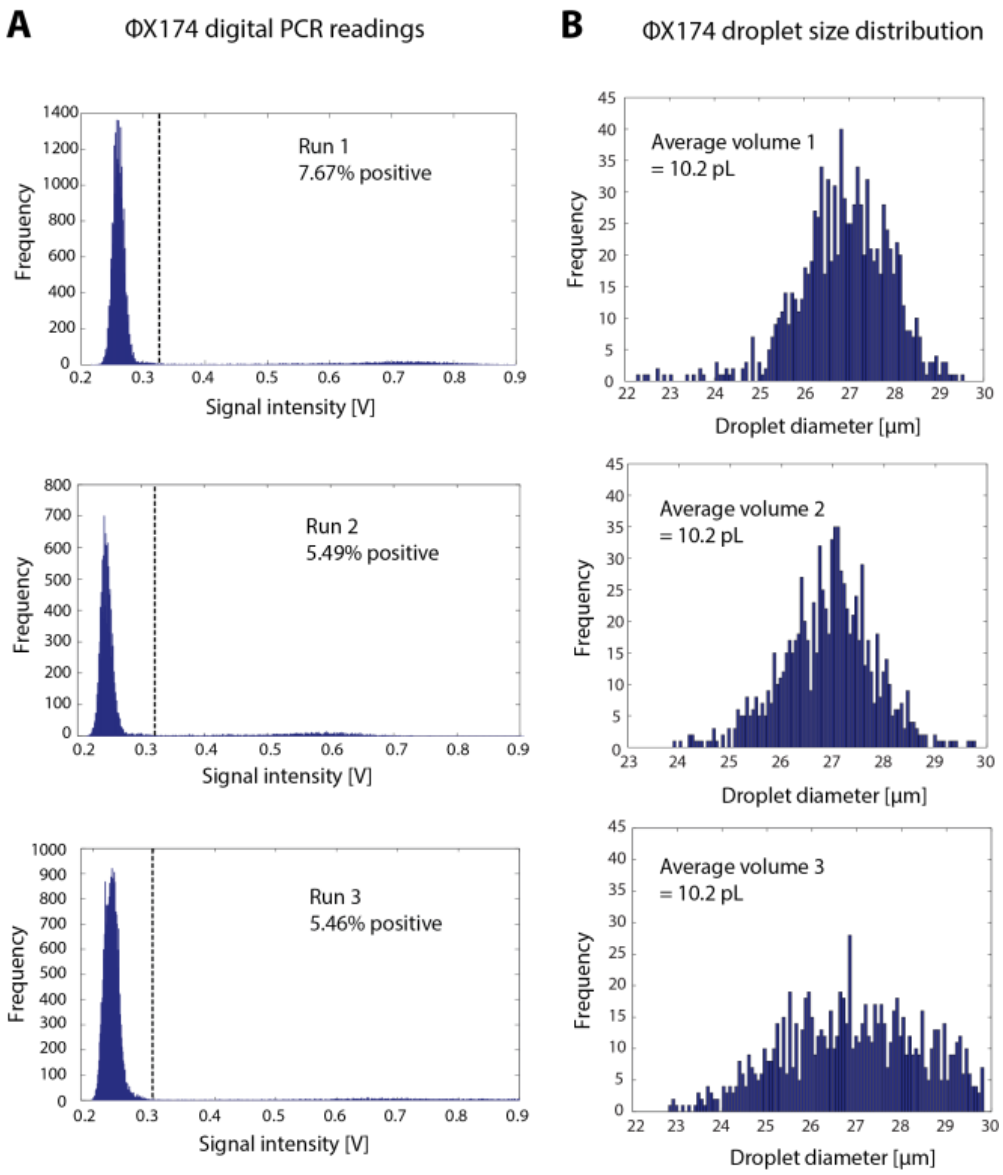
sequential droplet splitting, and using double emulsion FACS for the sorting[235], it should be possible to increase the throughput of the system by another order of magnitude, further increasing the number of cells that can be sorted for identifying rare virus-host relationships, such as identifying the host of chimeric RNA-DNA virus genomes recently discovered in numerous ecosystems[289,290]. PACS could also be used for screening novel potential pathogenic virus sequences for their hosts.

PACS and NGS form a particularly potent combination. PACS-sorted lysates can be recovered and subjected to deep sequencing of cellular DNA and/or RNA, enabling correlation of specific viruses with not only host cell genomes but also expression patterns. Virus-based PACS can be used, for example, to study how viruses modulate host cell gene expression, or if certain host genetic variants are more susceptible to virus infection. The ability to sort millions of entities based on nucleic acids is valuable for applications in microbiology, virology, and cell biology. In addition to sorting microbes based on infection by bacteriophage, PACS can be applied to mammalian cells infected with pathogens, such as HIV, malaria, or tuberculosis. By combining PACS with NGS, the presence of a specific pathogen can be correlated with host cell properties, such as somatic mutations or gene expression. Such investigations would be valuable for studying how different pathogens manipulate their hosts.

6.6 Supplemental Information

Primers for Taqman droplet PCR	**Sequence**
T4 probe	5'-/56-FAM/TCGCATTCT/ZEN/TCCTCTGATGGAGCA /3IABkFQ/ -3'
T4 FWD	5’-CCACAACTAACCGAGGAAGTAA-3’
T4 REV	5’-TGCGATATGCTATGGGTCTTG-3’
PhiX174 probe	5’-/56-FAM/ATG GAA CTG /ZEN/ACC AAA CGT CGT TAG GCC A/3IABkFQ/-3’
PhiX174 FWD	5’- GCGCTCTAATCTCTGGGCAT-3’
PhiX174 REV	5’- CAAAGAAACGCGGCACAGAA-3’
Lambda probe	5’- /56-FAM/AT ACT GAG C/ZEN/A CAT CAG CAG GAC GC/3IABkFQ/-3’
Lambda FWD	5’- GCC CTT CTT CAG GGC TTA AT-3’
Lambda REV	5’- CTC TGG CGG TGT TGA CAT AA-3’
Primers for qPCR	**Sequence**
E coli C FWD	5’-ACG CAG GGA TTT ACA GCA TAT AG-3’
E coli C REV	5’-GGG TGC TAT ATA ACG GTG TAC TG-3’
E coli K-12 FWD	5’-GACTACTTGAAGCTGTCCTTCC-3’
E coli K-12 REV	5’-CGCAGCTTCACCTTGTAGAT-3’

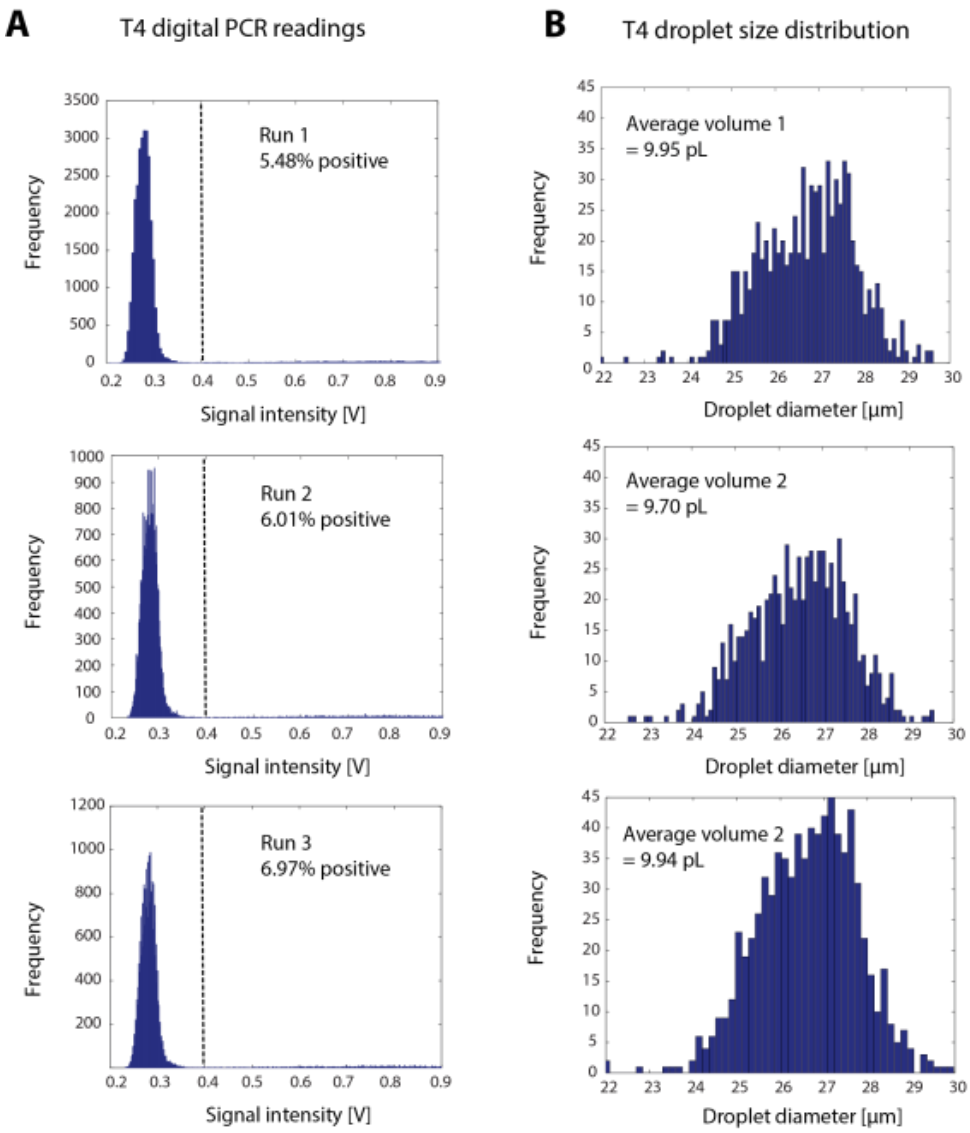
Table 6.1 Primers and TaqMan probe sequences



C

Run	Positive drops (%)	λ (particles/drop)	Virus conc. (/mL)
1	7.67	0.0798	7.82*10e6
2	5.49	0.0565	5.54*10e6
3	5.46	0.0561	5.35*10e6

Figure 6.5 Raw data measurements for ΦX174 virus. A) PMT measurements of 30,000 droplets. B) Droplet size measurements for 1,000 droplets. C) Virus concentration calculations using Poisson statistics for the 3 replicate measurements.



C

Run	Positive drops (%)	λ (particles/drop)	Virus conc. (/mL)
1	5.48	0.0798	8.02*10e6
2	6.01	0.0565	5.82*10e6
3	6.97	0.0722	7.26*10e6

Figure 6.6 Raw data measurements for T4 virus. A) PMT measurements of 30,000 droplets. B) Droplet size measurements for 1,000 droplets. C) Virus concentration calculations using Poisson statistics for the 3 replicate measurements.

Chapter 7: Optimization of PCR-Activated Cell Sorting (PACS) for robust and specific enrichment single cells

7.1 Introduction

Single cell analysis technologies are invaluable tools for quantifying and characterizing rare cells from heterogeneous populations. Single cell technologies could address fundamental questions in many areas of human health, from understanding cancer tumor mutation rates to characterizing the latent HIV reservoir. PCR-Activated Cell Sorting (PACS) enables researchers to detect and enrich complex, heterogeneous cell samples for specific cells of interest using PCR[280]. Eastburn et. al published the

In this chapter, we outline the steps involved in the PACS workflow. We then describe several optimization strategies that improve the reliability and robustness of PACS, which increases the accessibility of the technology to other researchers. Next, the application of PACS to measuring and characterizing the latent HIV reservoir is described. Finally, we suggest how the PACS workflow can be further optimized to enable greater utility and accessibility of the technology.

7.2 PACS Workflow

PCR-Activated Cell Sorting (PACS) is a new technology that detects a nucleic acid biomarker within single cells using PCR or RT-PCR, and then sorts cells based on this information[280]. In PACS, individual cells are encapsulated in droplets using microfluidics, lysed, and subjected to TaqMan PCR. TaqMan probes specific to the nucleic acid biomarker of interest produce detectable fluorescent signal in the droplet that can be used to trigger sorting. The sorted

cells are then available for downstream analysis such as qPCR or deep sequencing. The PACS workflow is illustrated in **Figure 7.1.**

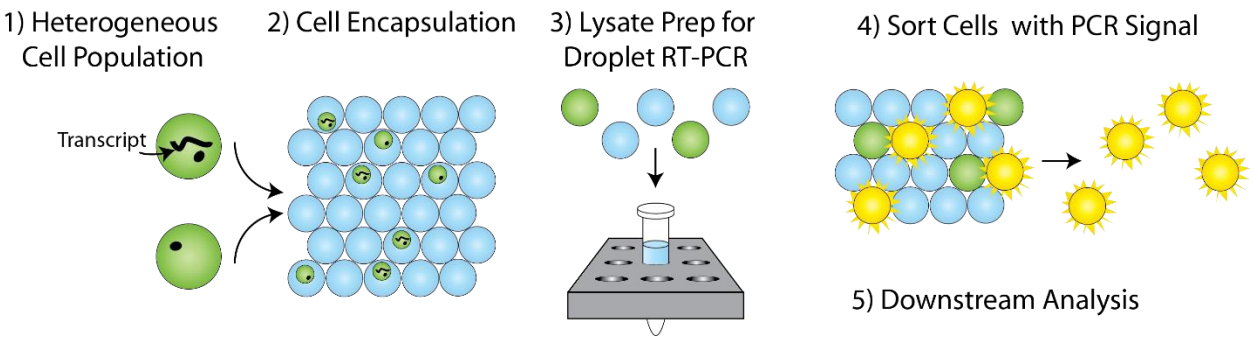


Figure 7.1 PCR-Activated Cell Sorting (PACS) workflow.

The PACS workflow can be conceptually divided into single cell RT-PCR and droplet sorting. The first step of the single cell RT-PCR reaction involves encapsulating fluorescently stained cells and cell lysis solution in monodisperse droplets, as depicted in Figure 7.2A. Cells are loaded into droplets such that, on average, one in every ten droplets contains a cell. The emulsion is then incubated off-chip to lyse the cells using Proteinase-K, which makes the cellular RNA accessible for interrogation by RT-PCR (Figure 7.2B). After the cells are lysed, the fluorescent cell stain fills the volume of the droplet (Figure 7.2C, 7.2D).

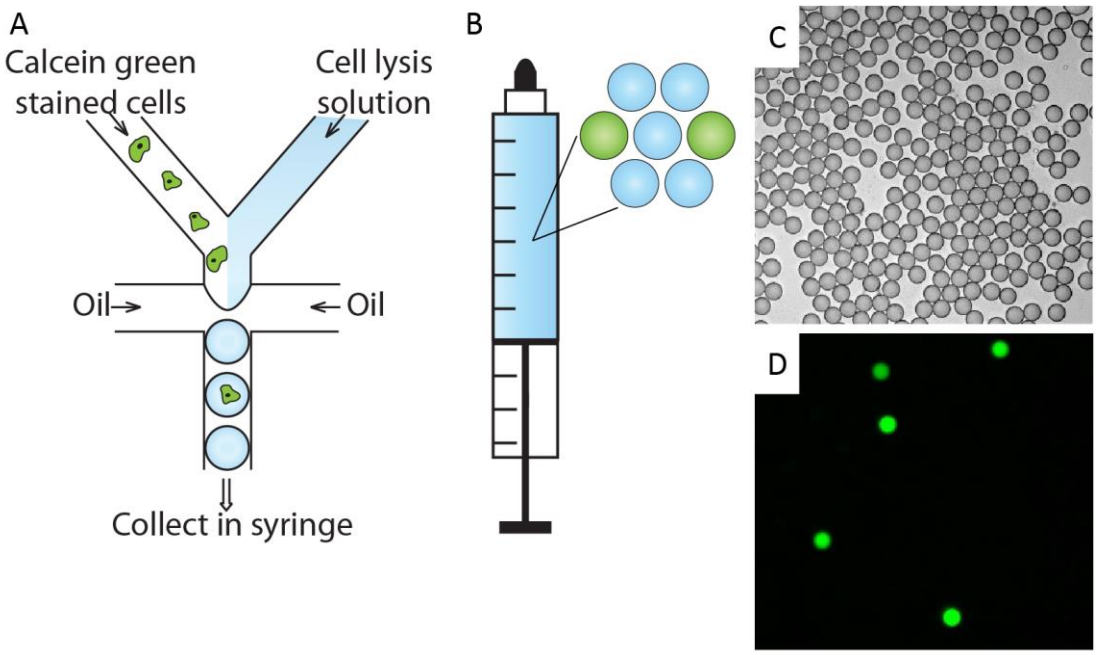


Figure 7.2 Cell encapsulation and lysis. A) Fluorescently stained cells and cell lysis solution are encapsulated into droplets. B) Droplets are incubated off-chip to lyse the cells. C) Bright field image of droplets after cell lysis. D) Fluorescent image of droplets after cell lysis.

The next step of the single-cell RT-PCR reaction involves adding RT-PCR reagents to the droplets containing cell lysate using the microfluidic device shown in Figure 7.3A. The re-injected emulsion travels through a filter that captures large, merged droplets while allowing monodisperse droplets to enter the device (Figure 7.3B). The cell lysate droplets are merged[286] with droplets containing RT-PCR reagents (Figure 7.3C) including TaqMan primers, probe, and RT-PCR mix. The merged droplets are mixed to evenly distribute the RT-PCR reagent and cell lysate throughout the volume of the droplet (Figure 7.3D). After mixing, the droplets are split into four droplets and collected for thermal cycling in a conventional benchtop thermal cycler (Figure 7.3E).

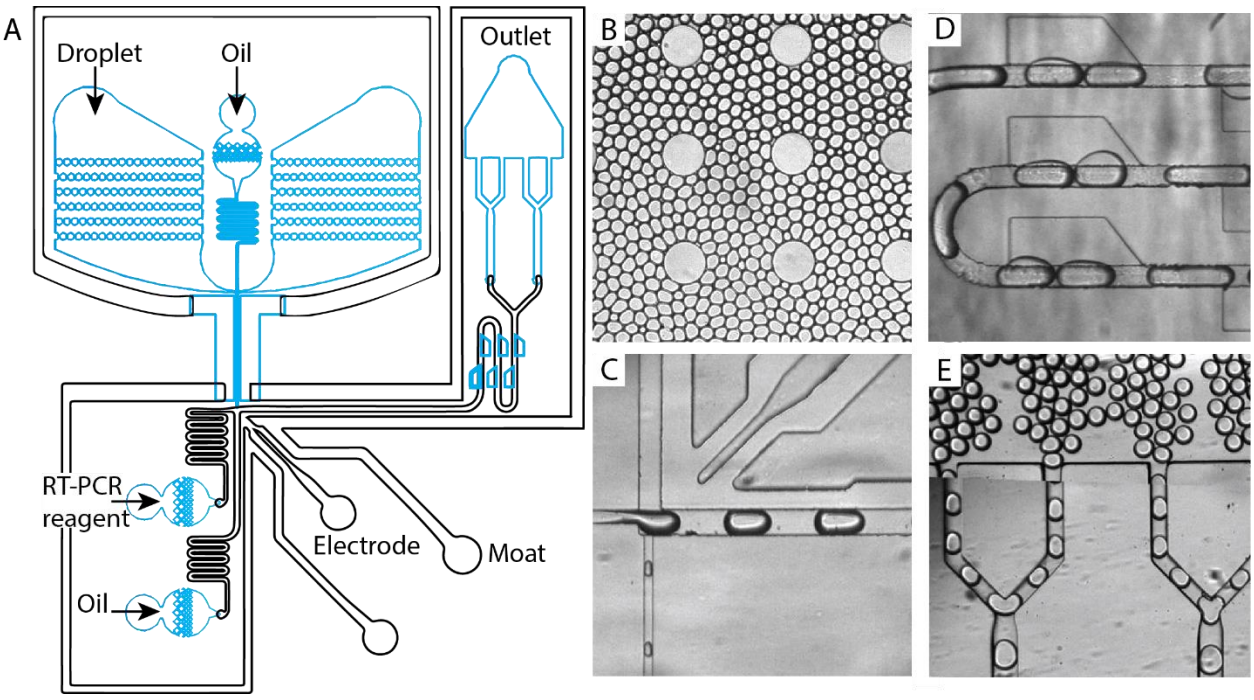


Figure 7.3 Addition of RT-PCR reagent to cell lysate drops. A) Schematic of microfluidic device for RT-PCR reagent addition. B) Bright field image of re-injected droplets. C) Bright field image of merging droplets containing cell lysate and droplets containing RT-PCR reagents. D) Bright field image of droplet mixing. E) Bright field image of droplet splitting.

After thermal cycling, the single cell RT-PCR reaction can be analyzed using fluorescence microscopy or microfluidic fluorescence detection[279]. Representative images of the single cell RT-PCR emulsion after thermal cycling are shown in Figure 7.4**.**

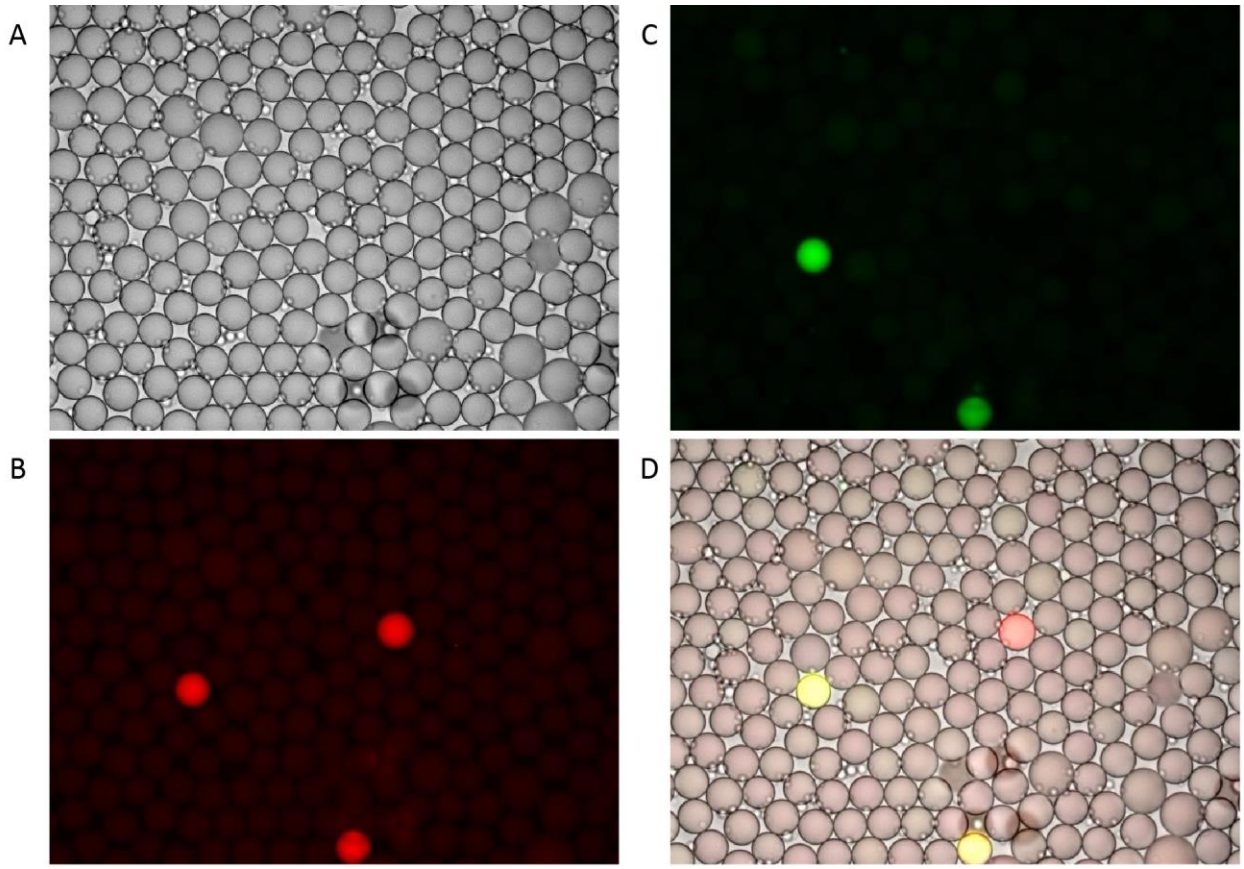


Figure 7.4 Single cell RT-PCR results. A) Bright field channel. B) FAM channel shows Calcein green cell stain. C) Cy5 channel shows Taqman amplification. D) Overlay of bright field, FAM, and Cy5 channels.

The second part of the PACS workflow involves droplet sorting. The single cell RT-PCR reaction produces droplets containing fluorescent signals for cell stain (Figure 7.4C) and TaqMan signal (Figure 7.4B) that is specific to a nucleic acid biomarker of interest. Droplets that are positive for both cell stain and TaqMan signal are detected and sorted using the microfluidic device shown in Figure 7.5. The thermal cycled droplets are re-injected and fluorescently excited using lasers that are focused upstream of the dielectrophoretic sorting junction[141,142,285]. The droplets emit fluorescent signal that is measured by PMTs. Threshold fluorescent values are designated in

a custom LabView program to designate droplets containing cell stain and TaqMan signal. When a drop has fluorescence intensity above the threshold level, the software sends a signal to the sorting electrode. The electrode applies a brief electric field to deflect the target droplets to a collection channel. All other droplets flow to a waste channel. The resulting emulsion is specifically enriched for the cell of interest and is suitable for downstream processing such as qPCR or sequencing.

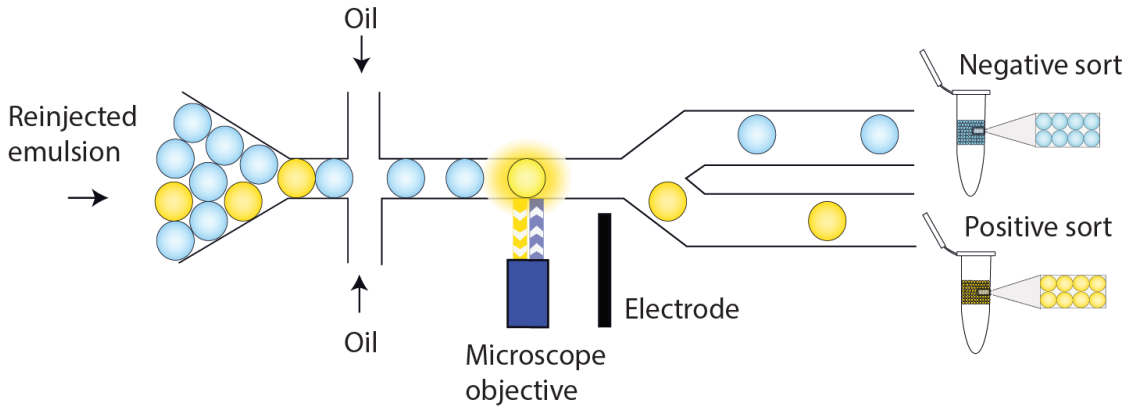


Figure 7.5 Microfluidic device for droplet detection and sorting.

7.3 Optimization of the PACS workflow

Eastburn et. al demonstrated PACS and its application to cancer cell sorting[280]. The published workflow is effective, but in practice it is difficult to duplicate the results because the method is susceptible to failure at several points along the workflow. However, it is essential for PACS to be robust in order to be accessible to other researchers and implemented accordingly. In order to improve the robustness of the workflow, optimization experiments are designed to minimize digital background, increase the signal to noise of the single cell RT-PCR reaction, and improve droplet stability.

Reducing digital background

The first microfluidic step in the PACS workflow involves encapsulating cells and cell lysis solution in droplets (Figure 7.2). Prior to droplet encapsulation, the cells are fluorescently stained, suspended in a density-matching solution to prevent settling, and loaded into a syringe that connects to the microfluidic device for droplet encapsulation. If a cell lyses before encapsulation, then it may release its genomic and transcriptional material into solution, and those sequences can be encapsulated with other, intact cells. The amplification of cell-free nucleotide sequences is called digital background, and digital background can contribute false positives to the PACS analysis. For example, if the oligonucleotide sequence that is the target of the single-cell RT-PCR reaction is encapsulated with a cell that does not contain the sequence, then PACS measures a false positive count. Therefore, it is important to minimize digital background in order to improve PACS accuracy.

An effective way to reduce digital background is to wash the cell sample immediately before encapsulation. To demonstrate this finding, single cell RT-PCR is performed on Calcein-stained 8E5 cells that are either washed immediately before encapsulation or allowed to rest for 30 minutes before encapsulation. After thermal cycling, the droplets are imaged using fluorescence microscopy and characterized as cell positive (Calcein+ TaqMan -), cell positive and TaqMan positive (Calcein+ TaqMan +), cell negative and TaqMan positive (Calcein- TaqMan +), or cell negative and TaqMan negative (Calcein- TaqMan -). The majority of droplets (~90%) are cell negative and TaqMan negative (Calcein- TaqMan -) and have no fluorescent signal. The fraction of droplets that have any positive fluorescent signal are shown in Figure 7.6. The cells that rested for 30 minutes before encapsulation are represented in Figure 7.6A and the cells that were washed immediately before encapsulation are in Figure 7.6B. Digital background is approximated using the fraction of Calcein- TaqMan + droplets, as these droplets contain cell-free nucleotide

sequences that could be co-encapsulated with a target-negative cell. The results show that the fraction of Calcein- TaqMan + droplets decreases when the cells are washed immediately before encapsulation. These results are consistent across two different TaqMan primer and probe combinations.

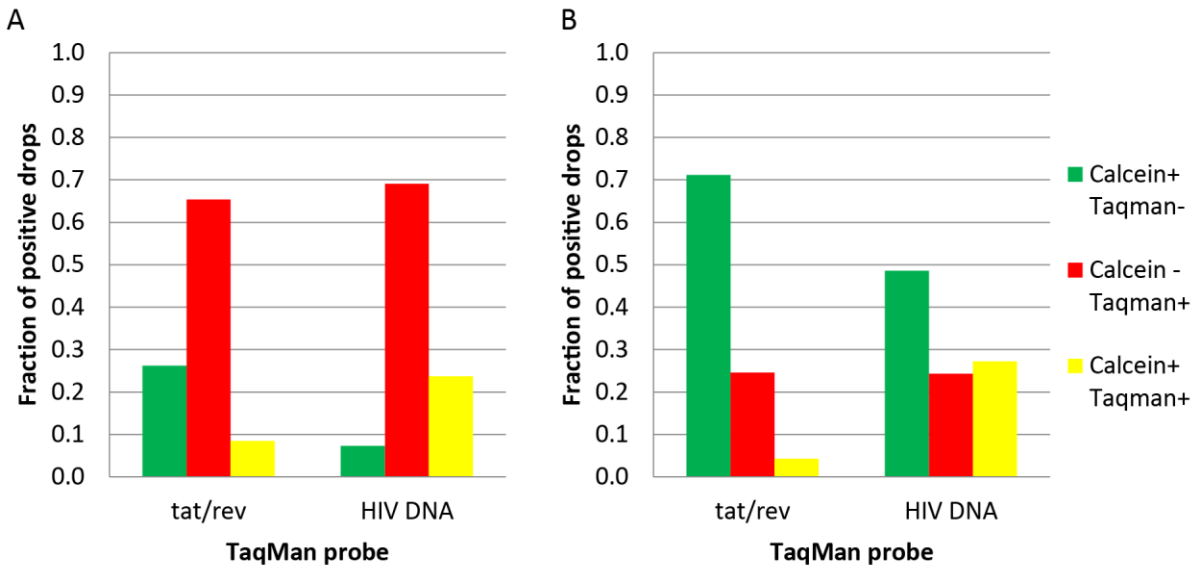


Figure 7.6 Fraction of positive drops for single cell RT-PCR performed on 8E5 cells with two different TaqMan primer and probe combinations. A) Cells rest for 30 minutes before encapsulation. B) Cells washed immediately before encapsulation.

Increasing signal to noise of single cell RT-PCR reaction

PACS uses fluorescence detection to trigger the sorting of single cells that contain a nucleic acid sequence of interest. The cells are fluorescently stained, and the TaqMan probes specific to the nucleic acid biomarker of interest produce detectable fluorescent signal that fills the volume of the droplet, as shown in Figure 7.4. Droplets are sorted if they contain the fluorescent signals for cell stain and TaqMan amplification. Therefore, it is important that the cell stain and TaqMan

probes generate strong fluorescent signal such that positive droplets can be confidently distinguished from negative droplets.

The cells are stained using Calcein dyes, which are cell-permeant dyes that generate fluorescent signal in viable cells when hydrolyzed by intracellular esterases. Calcein stains (Thermo Fischer) are available in violet (excitation/emission maxima 400/452 nm), green (excitation/emission maxima 495/515 nm), and red-orange (excitation/emission maxima 577/590 nm). To determine which Calcein stain yields the greatest signal to noise, single-cell RT-PCR is performed on 8E5 cells stained with Calcein violet, Calcein red-orange, or Calcein green. To stain the cells, Calcein is freshly reconstituted with DMSO to a final concentration of 2 µM, and 2 µL of stain is added for every 1 million cells. The cells incubate at room temperature in the dark for 30 minutes. After incubation, the cells are washed three times with PBS, resuspended in a density-matching solution, and loaded into a syringe for the single-cell RT-PCR workflow. After thermal cycling, the droplets are imaged using fluorescence microscopy and the signal to noise of the three Calcein stains is compared visually.

Figure 7 shows representative images of 8E5 cells stained with Calcein violet, Calcein red-orange, and Calcein green. The exposure and LUT settings are constant across the three images. The sample stained with Calcein green (Figure 7.7C) has a greater signal to noise ratio than the samples stained with Calcein violet (Figure 7.7A) and Calcein red-orange (Figure 7.7B).

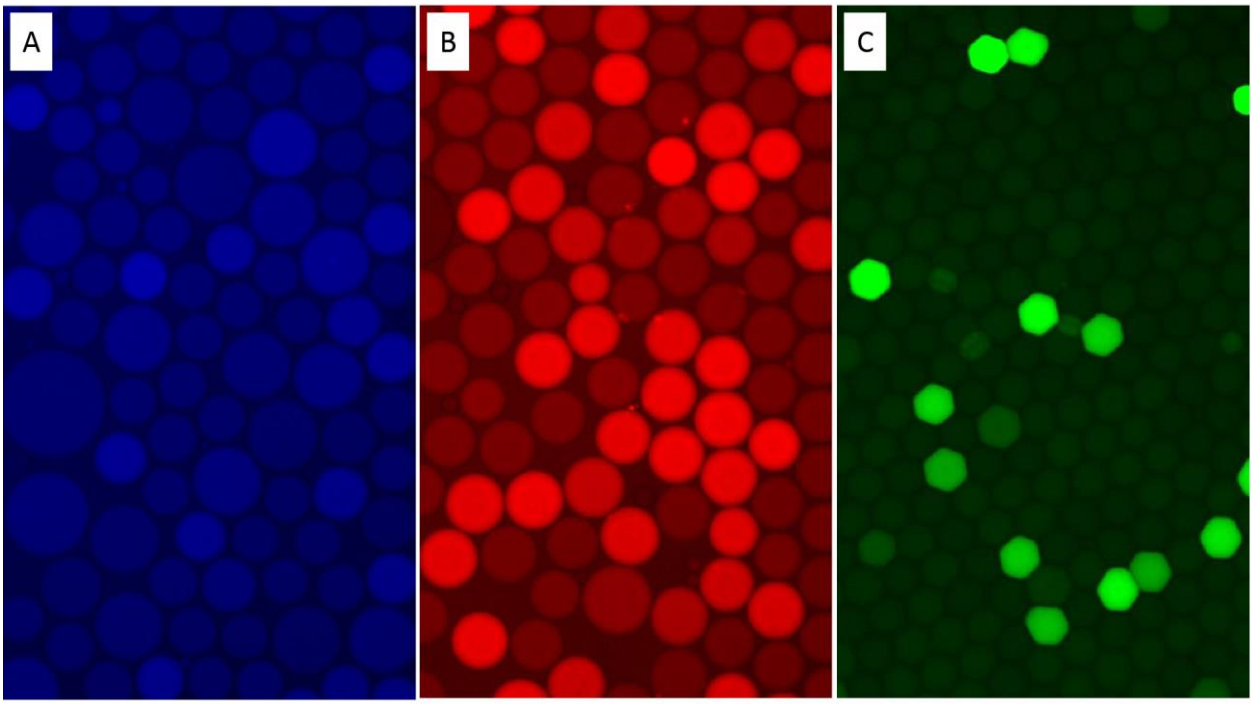


Figure 7.7 Single-cell RT-PCR is performed on 8E5 cells stained different Calcein dyes. A) Calcein violet. B) Calcein red-orange. C) Calcein green.

Further optimization testing indicates that Calcein green yields the best signal to noise when it is freshly reconstituted prior to use and when the emulsion is analyzed immediately upon completion of thermal cycling for the single cell RT-PCR reaction.

PACS sorting events are triggered by the presence of cell stain and TaqMan amplification in the same droplet, so successful sorting requires the cell stain and TaqMan probe to have strong, non-overlapping fluorescence spectra. TaqMan probes are available that occupy a wide range of excitation and emission spectra. However, for a given TaqMan fluorophore, the probe brightness can vary between reactions. Therefore, researchers using PACS should design multiple TaqMan primer and probe sets that are specific to nucleic acid biomarkers present in the cell type of interest and compare the fluorescence brightness to determine the best candidate set. Favorable TaqMan

candidates will have a binary signal with distinct separation between positive and negative droplets, such as the probe shown in Figure 7.8. The emulsion shows binary signal between positive and negative droplets (Figure 7.8A). A histogram of the mean intensity of the droplet fluorescence (Figure 7.8B) shows the positive droplets are at least 3x as bright as the negative droplets, so they can be readily gated to distinguish between positive and negative for sorting.

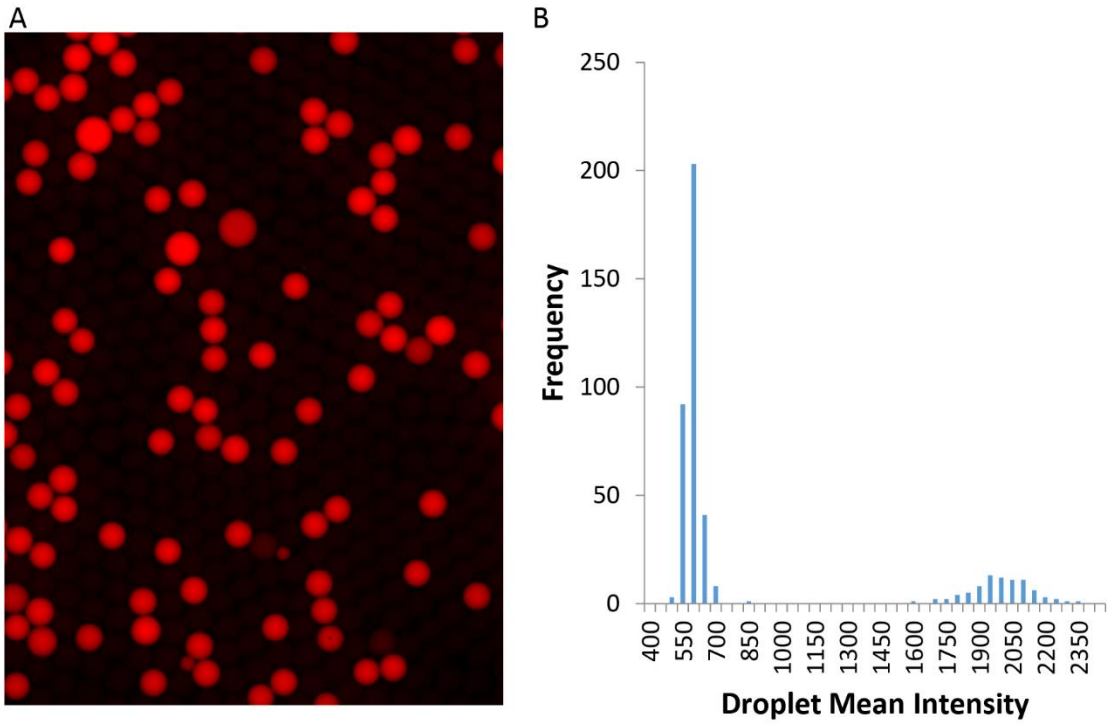


Figure 7.8 Digital droplet RT-PCR is used to evaluate signal of TaqMan probe candidates. A) Digital droplet RT-PCR using Cy5 Taqman probe. B) Histogram of droplet mean intensity for the digital droplet RT-PCR reaction.

Improving droplet stability

Droplet stability is critical for the reliability of PACS measurements, as uncontrolled droplet merger compromises direct quantitative measurements and downstream analyses. PACS

employs surfactants to prevent droplet coalescence by reducing surface tension, yet heat, extraneous movement, and static can cause droplet merger even when surfactants are used.

Several steps were taken to improve droplet stability in the PACS workflow. First, the cell lysis solution used in the cell encapsulation step was optimized from the original buffer used in Easburn et. al[280]. The compositions of the original and optimized buffers are detailed in Table 7.1 and Table 7.2, respectively. Figure 7.9 shows representative images of single cell RT-PCR results for cells that are encapsulated with the original buffer (Figure 7.9A) and the optimized buffer (Figure 7.9B). The coefficient of variation in droplet volume decreased from 0.49 to 0.09, where a CV of Figure 7.9B 0.0 would indicate a perfectly monodisperse emulsion. Importantly, the optimized encapsulation buffer is compatible with the single cell RT-PCR reaction.

After the cells are encapsulated with cell lysis solution, the emulsion is heated off-chip to 55° C for 30 minutes and then 95° C for 10 minutes to activate and deactivate Proteinase-K, respectively. After heating, the emulsion is re-injected into a second microfluidic device that merges cell lysate droplets with RT-PCR reagent (Figure 7.3). The process of heating, handling, and re-injecting the emulsion can introduce polydispersity to the sample. These concerns motivated the optimization of the workflow configuration. Eastburn et. al batch collected the emulsion from the cell encapsulation step (Figure 7.2) into PCR tubes and used a thermal cycler to heat the sample for Proteinase-K treatment[280]. Thermal cyclers offer accurate and precise temperature control and timing of the heating steps, but droplet stability is compromised when the emulsion is transferred from the PCR tube into a syringe for sample re-injection. An alternative method was developed that reduces droplet handling without compromising the precision and accuracy of the heating steps. First, the entire emulsion from the cell encapsulation step is collected into a new syringe. The syringe is capped and completely submerged in water heat baths that

consist of 12-quart polycarbonate containers and sous vide controllers (Anova) programmed to 55° C and 95° C. The syringe is submerged for 30 minutes in the 55° C water heat bath and then transferred to the 95° C for 10 minutes. After incubation, the syringe cap is replaced with a needle, thus preparing the sample for re-injection to the second microfluidic device. The sous vide controllers provide automated temperature control with .01° C accuracy, and incubating the emulsion in the re-injection syringe ensures that the sample is handled minimally. This improved heat bath configuration is more robust to user error been shown to result is less droplet polydispersity than the published PACS workflow[280].

Volume (uL)	Component
800	125 mM Tris buffer pH 8.0
200	Proteinase K

Table 7.1 Cell lysis solution from Eastburn et. al

Volume (uL)	Component
125	125 mM Tris buffer pH 8.0
170	50% PEG
10	5 mM EDTA
40	50% Tween-20
555	DI water
100	Proteinase K

Table 7.2 Optimized cell lysis solution.

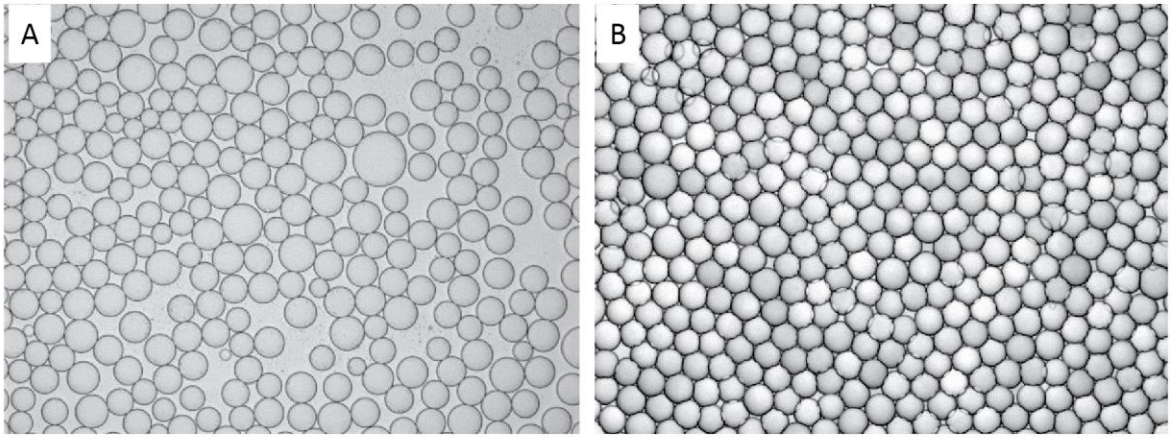


Figure 7.9 Bright field images of single cell RT-PCR samples. A) Cells encapsulated with the original cell lysis solution. B) Cells encapsulated with the optimized cell lysis solution.

7.4 PACS for enrichment of latent HIV reservoir

One exciting application of PACS is the measurement and enrichment of the latent HIV reservoir. The ability to quantify the latent reservoir would enable researchers to evaluate the efficacy of new therapeutic interventions, and it would have potential clinical impact for monitoring patient health. PACS can selectively enrich patient samples for latently infected cells, and allows for detailed characterization of latently infected cells with downstream analyses.

Background

Antiretroviral Therapy (ART) inhibits viral replication and improves the quality of life of many HIV-infected patients. The initiation of ART results in a rapid drop in plasma viral load and in a substantial reduction in the number of cells carrying proviral DNA in both blood and tissues[291]. However, ART does not eradicate HIV, and the spread of the virus almost immediately resumes upon cessation of ART in all but exceptional cases[291]. ART does not eradicate HIV due to the persistence of latently infected cells[292,293]. Latently infected cells contain integrated HIV DNA that are replication competent but transcriptionally silent[293]. The latent reservoir is considered a major hurdle to developing a cure for HIV, and interventions are being developed to reduce the size of the latent reservoir or eliminate it altogether. In order to assess the effectiveness any given intervention, researchers need tools capable of quantifying and characterizing the latent reservoir in response to the intervention. Quantifying the latent reservoir is challenging due to the relatively low frequency of cells in patients on ART and the fact that among all HIV DNA positive cells, only a fraction of viruses are replication competent[294].

Current tools are inadequate for quantifying and characterizing the latent reservoir. The Viral Outgrowth Assay (VOA) is the current gold standard for quantifying the latent reservoir.

This assay provides a functional measure of the latent reservoir, as it quantifies the number of cells that can release virus capable of replication when activated[294]. However, VOA is relatively insensitive, requiring 10^4 to 10^6 virions for a reliable measurement[295]. VOA is also time and labor intensive, requiring 3-4 weeks to complete. Another technology that offers improved sensitivity is the Fluidigm C1 and Biomark system, which interrogates 96 single cells for expression of selected genes. While this system is robust, the throughput is too low to identify and characterize the rare HIV latent cells. A technology that provides improved throughput with single-cell sensitivity is fluorescence in-situ hybridization with fluorescence-activated cell sorting (FISH-FACS), as described in Chapter 6. However, this technique requires cell fixation, which can degrade nucleic acids and introduce mutations that compromise downstream sequencing data. PACS is well suited for quantifying and characterizing the latent HIV reservoir. It does not require cell fixation, and it provides single-cell sensitivity at higher throughput than the Fluidigm C1 system.

Results

In order to identify cells comprising the functional HIV reservoir, the single cell RT-PCR reaction can target a previously published TaqMan RT-PCR assay for the multiply spliced (ms) RNA species that encodes the Tat and Rev proteins of HIV[292,296,297] (Table 7.3). Since Tat mediates high level viral transcription of the proviral DNA and Rev is responsible for efficient translocation of a majority of viral mRNA species into the cytoplasm for translation, expression of msRNA encoding *tat/rev* may be viewed as a surrogate for the efficient expression of viral proteins[297,298].

To demonstrate the utility of the tat/rev TaqMan assay for detection of HIV-infected cells, single cell RT-PCR is performed on 8E5 cells. The 8E5 clonal cell line is a lymphoblast HIV positive cell line that contains a single integrated copy of defective HIV provirus[299]. The results of

the single cell RT-PCR reaction using the optimized workflow techniques are shown in Figure 7.4. The images demonstrate that the TaqMan probe and Calcein cell stain have favorable signal to noise, and the emulsion remains acceptably monodisperse throughout the single cell RT-PCR reaction.

tat/rev forward primer	CTTAGGCATCTCCTATGGCAGGAA
tat/rev reverse primer	GGATCTGTCTCTGTCTCTCTCTCCACC
tat/rev TaqMan probe	Cy5-AGGGGACCCGACAGGCCC-BHQ2

Table 7.3 PCR primers and probe for tat/rev

After validating the TaqMan reaction, it is necessary to confirm that Calcein green is effective at staining patient cells. First, peripheral blood is collected from patients, Ficoll gradient centrifugation is used to isolate the peripheral blood mononuclear cells (PBMCs), and negative selection is used to isolate the CD4+ T cells[300]. The cells are activated by stimulating with for 24 hours. To stain the activated CD4+ T cells, Calcein green is freshly reconstituted with DMSO to a final concentration of 2 μM, and 2 μL of stain is added for every 1 million cells. The cells incubate at room temperature in the dark for 30 minutes. After incubation, the cells are washed three times with PBS, resuspended in a density-matching solution of PBS with 17% Optiprep. The sample is then loaded into a syringe for encapsulation with cell lysis solution (Table 7.2) and incubated off-chip to allow cell lysis to occur. The lysed cells are imaged using fluorescence microscopy. The sample has good signal to noise (Figure 7.10A), and a histogram of the mean intensity of the droplet fluorescence (Figure 7.10B) illustrates how the sample can be readily gated to distinguish between positive and negative for sorting.

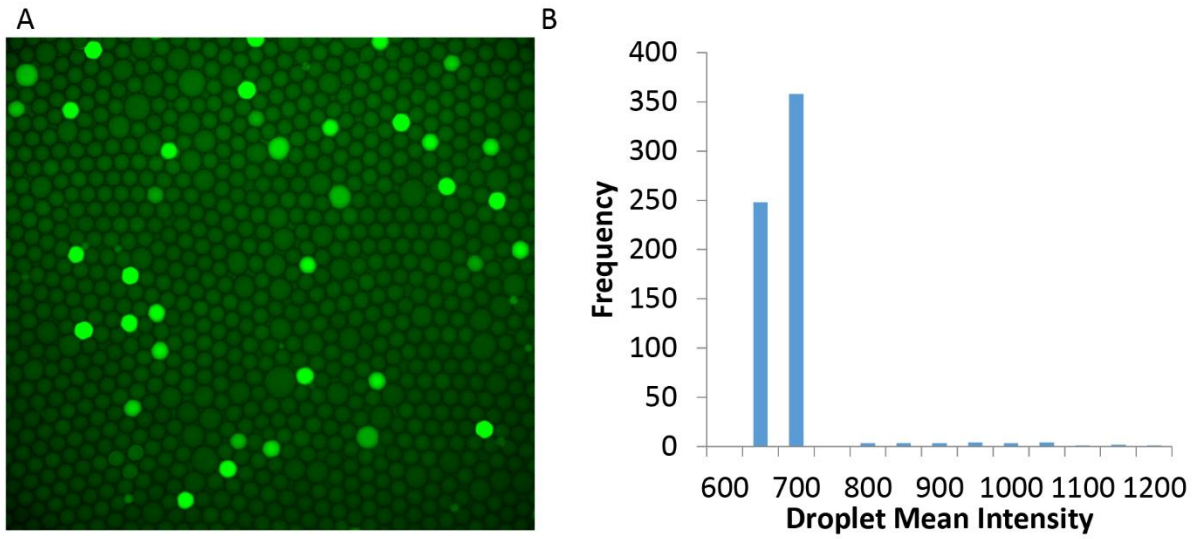


Figure 7.10 Calcein green stained patient CD4+ T cells. A) Fluorescent image of patient cells CD4+ T cells that are stained, encapsulated, and lysed. B) Histogram of droplet mean intensity of the Calcein green stained patient CD4+ T cells.

Discussion

Further optimization of the throughput of the PACS workflow is necessary before it can be applied to study HIV latency. First, the throughput of the existing PACS workflow precludes its application to quantifying the latent HIV reservoir. Eastburn et. al used PACS to perform single cell RT-PCR and sorting on 132,000 single cells, which represents a 500-fold improvement over other methods[280]. While the frequency of latently infected cells in a patient sample is highly variable, it is expected that, at most, 1/2000 to 1/200,000 CD4 T cells in blood are HIV+. If the patient sample has 1/200,000 CD4 T cells, then the current PACS workflow would optimistically be able to detect and enrich 1-10 cells. The existing workflow is suited for applications that require the quantification and enrichment of cell types that are less rare, and an optimized cell loading strategy is necessary for application to HIV latency.

After the PACS workflow is optimized to handle higher throughput, the single RT-PCR assay should be benchmarked against existing methods. Measurements made using the proposed technique can be compared to values obtained using VOA and digital PCR for HIV DNA. The frequency of HIV+ cells measured using single cell RT-PCR assay is expected be greater than that of the VOA, which has been reported to underestimate the size of the latent reservoir by 60-fold[301]. Furthermore, the frequency of HIV+ cells measured using the proposed technique is expected to be less than that of digital PCR for HIV DNA. Previous groups have used digital PCR on purified DNA from CD4+T cells to detect latently infected cells. On average, this method estimates the latent reservoir to be two orders of magnitude greater than that of VOA, as it captures both replication-competent and replication-incompetent viruses in its count[294,302]. As the proposed technique detects the *tat/rev* transcript, it is expected it will provide a more reliable measurement of the functional HIV reservoir.

The ability to characterize an enriched pool of latently infected cells using deep sequencing holds particular promise for revealing cellular mechanisms responsible for HIV latency and the effects of pharmacologic anti-latency therapies. However, research outside the scope of this dissertation reveals that cellular RNA degrades during the single cell RT-PCR workflow, particularly during the incubation step for Proteinase-K inactivation. Such degradation compromises the ability to characterize the enriched latently infected cells using deep sequencing. Further optimization of the PACS workflow is required to preserve the integrity of cellular transcriptional sequences.

PACS provides a high-throughput, quantitative measure of the functional HIV reservoir, which has potential clinical impact for monitoring patient health. The ability to enrich for HIV+ patient cells is and downstream sequencing analysis. Gene expression analysis of HIV+ cells could

enable the study of cellular mechanisms governing latency and the effects of new anti-latency therapies. Furthermore, the RT-PCR assay could be modified to enable single-cell gene expression analysis, which could be used to investigate the heterogeneity of latently infected cells in various tissues and blood.

References

1. Wang, W. *et al.* Identification of small-molecule inducers of pancreatic beta-cell expansion. *Proc. Natl. Acad. Sci. U. S. A.* **106,** 1427–1432 (2009).
2. Burbaum, J. J. Miniaturization technologies in HTS: How fast, how small, how soon? *Drug Discov. Today* **3,** 313–322 (1998).
3. Major, J. Challenges and Opportunities in High Throughput Screening: Implications for New Technologies. *J. Biomol. Screen.* **3,** 13–17 (1998).
4. Aharoni, A., Griffiths, A. D. & Tawfik, D. S. High-throughput screens and selections of enzyme-encoding genes. *Curr. Opin. Chem. Biol.* **9,** 210–216 (2005).
5. Dalby, P. A. Optimising enzyme function by directed evolution. *Curr. Opin. Struct. Biol.* **13,** 500–505 (2003).
6. Wittrup, K. D. Protein engineering by cell-surface display. *Curr. Opin. Biotechnol.* **12,** 395–399 (2001).
7. Liebeton, K. *et al.* Directed evolution of an enantioselective lipase. *Chem. Biol.* **7,** 709–718 (2000).
8. Daugherty, P. S., Iverson, B. L. & Georgiou, G. Flow cytometric screening of cell-based libraries. *J. Immunol. Methods* **243,** 211–227 (2000).
9. Aebersold, R. & Mann, M. Mass spectrometry-based proteomics. *Nature* **422,** 198–207 (2003).
10. Boder, E. T. & Wittrup, K. D. Yeast surface display for screening combinatorial polypeptide libraries. *Nat. Biotechnol.* **15,** 553–557 (1997).
11. Fernandez-Gacio, A., Uguen, M. & Fastrez, J. Phage display as a tool for the directed evolution of enzymes. *Trends Biotechnol.* **21,** 408–414 (2003).

12. Forrer, P., Jung, S. & Pluckthun, A. Beyond binding: using phage display to select for structure, folding and enzymatic activity in proteins. *Curr.Opin.Struct.Biol.* **9,** 514–520 (1999).

13. Amstutz, P., Forrer, P., Zahnd, C. & Pl??ckthun, A. In vitro display technologies: Novel developments and applications. *Curr. Opin. Biotechnol.* **12,** 400–405 (2001).

14. Glieder, A., Farinas, E. T. & Arnold, F. H. Laboratory evolution of a soluble, self-sufficient, highly active alkane hydroxylase. *Nat. Biotechnol.* **20,** 1135–1139 (2002).

15. Jaeger, K. E. & Eggert, T. Lipases for biotechnology. *Curr. Opin. Biotechnol.* **13,** 390–397 (2002).

16. Turner, N. J. Directed evolution of enzymes for applied biocatalysis. *Trends Biotechnol.* **21,** 474–478 (2003).

17. Williams, G. J., Zhang, C. & Thorson, J. S. Expanding the promiscuity of a natural-product glycosyltransferase by directed evolution. *Nat. Chem. Biol.* **3,** 657–662 (2007).

18. Otten, L. G., Schaffer, M. L., Villiers, B. R. M., Stachelhaus, T. & Hollfelder, F. An optimized ATP/PPi-exchange assay in 96-well format for screening of adenylation domains for applications in combinatorial biosynthesis. *Biotechnol. J.* **2,** 232–240 (2007).

19. Polizzi, K. M. *et al.* Pooling for improved screening of combinatorial libraries for directed evolution. *Biotechnol. Prog.* **22,** 961–967 (2006).

20. Tee, K. L. & Schwaneberg, U. A screening system for the directed evolution of epoxygenases: Importance of position 184 in P450 BM3 for stereoselective styrene epoxidation. *Angew. Chemie - Int. Ed.* **45,** 5380–5383 (2006).

21. Teh, S.-Y., Lin, R., Hung, L.-H. & Lee, A. P. Droplet microfluidics. *Lab Chip* **8,** 198–220 (2008).

22. Rajendrani Mukhopadhyay. When PDMS isn't the best. *Anal. Chem.* 1–5 (2007). doi:10.1007/s13398-014-0173-7.2

23. Zeng, Y., Novak, R., Shuga, J., Smith, M. T. & Mathies, R. A. High-performance single cell genetic analysis using microfluidic emulsion generator arrays. *Anal. Chem.* **82,** 3183–3190 (2010).

24. Wu, N. *et al.* A PMMA microfluidic droplet platform for in vitro protein expression using crude E. coli S30 extract. *Lab Chip* **9,** 3391–3398 (2009).

25. Lorenz, H. *et al.* SU-8: a low-cost negative resist for MEMS. *J. Micromechanics Microengineering* **7,** 121–124 (1997).

26. McDonald, J. C. *et al.* Fabrication of microfluidic systems in poly(dimethylsiloxane). *Electrophoresis* **21,** 27–40 (2000).

27. Rotem, A., Abate, A. R., Utada, A. S., Van Steijn, V. & Weitz, D. a. Drop formation in non-planar microfluidic devices. *Lab Chip* **12,** 4263–4268 (2012).

28. Haubert, K., Drier, T. & Beebe, D. PDMS bonding by means of a portable, low-cost corona system. *Lab Chip* **6,** 1548–9 (2006).

29. Fritz, J. L. & Owen, M. J. Hydrophobic Recovery of Plasma-Treated Polydimethylsiloxane. *J. Adhes.* **54,** 33–45 (1995).

30. Dangla, R., Gallaire, F. & Baroud, C. N. Microchannel deformations due to solvent-induced PDMS swelling. *Lab Chip* **10,** 2972–2978 (2010).

31. He, M. *et al.* Selective encapsulation of single cells and subcellular organelles into picoliter- and femtoliter-volume droplets. *Anal. Chem.* **77,** 1539–1544 (2005).

32. Lee, J. N., Park, C. & Whitesides, G. M. Solvent Compatibility of Poly(dimethylsiloxane)-Based Microfluidic Devices. *Anal. Chem.* **75,** 6544–6554 (2003).

33. Courtois, F. *et al.* An integrated device for monitoring time-dependent in vitro expression from single genes in picolitre droplets. *ChemBioChem* **9,** 439–446 (2008).

34. Kumaresan, P., Yang, C. J., Cronier, S. A., Blazej, R. G. & Mathies, R. A. High-throughput single copy DNA amplification and cell analysis in engineered nanoliter droplets. *Anal. Chem.* **80,** 3522–3529 (2008).

35. Chen, F. *et al.* Chemical transfection of cells in picoliter aqueous droplets in fluorocarbon oil. *Anal. Chem.* **83,** 8816–8820 (2011).

36. Studer, a *et al.* Fluorous synthesis: a fluorous-phase strategy for improving separation efficiency in organic synthesis. *Science* **275,** 823–826 (1997).

37. Curran, D. P. Strategy-level separations in organic synthesis: From planning to practice (vol 37, pg 1174, 1998). *Angew. Chemie-International Ed.* **37,** 2292 (1998).

38. Lowe, K. C., Davey, M. R. & Power, J. B. Perfluorochemicals: Their applications and benefits to cell culture. *Trends Biotechnol.* **16,** 272–278 (1998).

39. Schmitz, C. H. J., Rowat, A. C., Köster, S. & Weitz, D. A. Dropspots: a picoliter array in a microfluidic device. *Lab Chip* **9,** 44–49 (2009).

40. Brouzes, E. *et al.* Droplet microfluidic technology for single-cell high-throughput screening. *Proc. Natl. Acad. Sci. U. S. A.* **106,** 14195–200 (2009).

41. Pan, J. *et al.* Quantitative tracking of the growth of individual algal cells in microdroplet compartments. *Integr. Biol.* **3,** 1043–1051 (2011).

42. Clausell-Tormos, J. *et al.* Droplet-Based Microfluidic Platforms for the Encapsulation and Screening of Mammalian Cells and Multicellular Organisms. *Chem. Biol.* **15,** 427–437 (2008).

43. Sadtler, V. M., Krafft, M. P. & Riess, J. G. Achieving stable, reverse water-in-

fluorocarbon emulsions. *Angew. Chemie* **35,** 1976–1978 (1996).

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

45. Bai, Y. *et al.* A double droplet trap system for studying mass transport across a droplet-droplet interface. *Lab Chip* **10,** 1281–1285 (2010).

46. Baret, J.-C. Surfactants in droplet-based microfluidics. *Lab Chip* **12,** 422–433 (2012).

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

48. Roach, L. S., Song, H. & Ismagilov, R. F. Controlling nonspecific protein adsorption in a plug-based microfluidic system by controlling interfacial chemistry using fluorous-phase surfactants. *Anal. Chem.* **77,** 785–796 (2005).

49. Sadtler, V. M., Jeanneaux, F., Pierre Krafft, M., Rábai, J. & Riess, J. G. Perfluoroalkylated amphiphiles with a morpholinophosphate or a dimorpholinophosphate polar head group. *New J. Chem.* **22,** 609–613 (1998).

50. Liau, A. *et al.* Mixing crowded biological solutions in milliseconds. *Anal. Chem.* **77,** 7618–7625 (2005).

51. Courtois, F. *et al.* Controlling the retention of small molecules in emulsion microdroplets for use in cell-based assays. *Anal. Chem.* **81,** 3008–3016 (2009).

52. Kaltenbach, M., Devenish, S. R. a. & Hollfelder, F. A simple method to evaluate the biochemical compatibility of oil/surfactant mixtures for experiments in microdroplets. *Lab Chip* **12,** 4185 (2012).

53. Paegel, B. M. & Joyce, G. F. Microfluidic compartmentalized directed evolution. *Chem. Biol.* **17,** 717–724 (2010).

54. Christopher, G. F. & Anna, S. L. Microfluidic methods for generating continuous droplet streams. *J. Phys. D. Appl. Phys.* **40,** R319–R336 (2007).

55. Vladisavljević, G. T., Kobayashi, I. & Nakajima, M. Production of uniform droplets using membrane, microchannel and microfluidic emulsification devices. *Microfluid. Nanofluidics* **13,** 151–178 (2012).

56. Seemann, R., Brinkmann, M., Pfohl, T. & Herminghaus, S. Droplet based microfluidics. *Rep. Prog. Phys.* **75,** 16601 (2012).

57. Utada, A. S., Fernandez-Nieves, A., Stone, H. A. & Weitz, D. A. Dripping to jetting transitions in coflowing liquid streams. *Phys. Rev. Lett.* **99,** 1–4 (2007).

58. Utada, A. S., Fernandez-Nieves, A., Gordillo, J. M. & Weitz, D. A. Absolute instability of a liquid jet in a coflowing stream. *Phys. Rev. Lett.* **100,** 1–4 (2008).

59. Cramer, C., Fischer, P. & Windhab, E. J. Drop formation in a co-flowing ambient fluid. *Chem. Eng. Sci.* **59,** 3045–3058 (2004).

60. Guillot, P., Colin, A. & Ajdari, A. Stability of a jet in confined pressure-driven biphasic flows at low Reynolds number in various geometries. *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.* **78,** (2008).

61. Thorsen, T., Roberts, R. W., Arnold, F. H. & Quake, S. R. Dynamic pattern formation in a vesicle-generating microfluidic device. *Phys. Rev. Lett.* **86,** 4163–4166 (2001).

62. Garstecki, P., Fuerstman, M. J., Stone, H. a & Whitesides, G. M. Formation of droplets and bubbles in a microfluidic T-junction-scaling and mechanism of break-up. *Lab Chip* **6,** 437–446 (2006).

63. Anna, S. L., Bontoux, N. & Stone, H. A. Formation of dispersions using ‘flow focusing’ in microchannels. *Appl. Phys. Lett.* **82,** 364–366 (2003).

64. Abate, A. R. *et al.* Impact of inlet channel geometry on microfluidic drop formation. *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.* **80,** 1–5 (2009).

65. De Menech, M., Garstecki, P., Jousse, F. & Stone, H. a. Transition from squeezing to dripping in a microfluidic T-shaped junction. *J. Fluid Mech.* **595,** 141–161 (2008).

66. Nie, Z. *et al.* Emulsification in a microfluidic flow-focusing device: Effect of the viscosities of the liquids. *Microfluid. Nanofluidics* **5,** 585–594 (2008).

67. Link, D. R., Anna, S. L., Weitz, D. a & Stone, H. a. Geometrically mediated breakup of drops in microfluidic devices. *Phys. Rev. Lett.* **92,** 54503 (2004).

68. Abate, A. R. & Weitz, D. a. Faster multiple emulsification with drop splitting. *Lab Chip* **11,** 1911–5 (2011).

69. Abate, A. R. & Weitz, D. a. Air-bubble-triggered drop formation in microfluidics. *Lab Chip* **11,** 1713–1716 (2011).

70. Nisisako, T., Torii, T., Takahashi, T. & Takizawa, Y. Synthesis of Monodisperse Bicolored Janus Particles with Electrical Anisotropy Using a Microfluidic Co-Flow System. *Adv. Mater.* **18,** 1152–1156 (2006).

71. Nisisako, T. & Torii, T. Microfluidic large-scale integration on a chip for mass production of monodisperse droplets and particles. *Lab Chip* **8,** 287–293 (2008).

72. Abate, A. R., Romanowsky, M. B., Agresti, J. J. & Weitz, D. A. Valve-based flow focusing for drop formation. *Appl. Phys. Lett.* **94,** (2009).

73. Priest, C., Herminghaus, S. & Seemann, R. Generation of monodisperse gel emulsions in a microfluidic device. *Appl. Phys. Lett.* **88,** 24106 (2006).

74. Chokkalingam, V., Herminghaus, S. & Seemann, R. Self-synchronizing pairwise production of monodisperse droplets by microfluidic step emulsification. *Appl. Phys. Lett.*

93, 10–13 (2008).

75. Köster, S. *et al.* Drop-based microfluidic devices for encapsulation of single cells. *Lab Chip* **8,** 1110–1115 (2008).

76. Huebner, A. *et al.* Quantitative detection of protein expression in single cells using droplet microfluidics. *Chem. Commun.* **2,** 1218–1220 (2007).

77. Kemna, E. W. M. *et al.* High-yield cell ordering and deterministic cell-in-droplet encapsulation using Dean flow in a curved microchannel. *Lab Chip* **12,** 2881–2887 (2012).

78. 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).

79. Chabert, M. & Viovy, J.-L. Microfluidic high-throughput encapsulation and hydrodynamic self-sorting of single cells. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 3191–3196 (2008).

80. Um, E., Lee, S.-G. G. & Park, J.-K. K. Random breakup of microdroplets for single-cell encapsulation. *Appl. Phys. Lett.* **97,** 153703 (2010).

81. Tawfik, D. S. & Griffiths, A. D. Man-made cell-like compartments for molecular evolution. *Nat. Biotechnol.* **16,** 652–656 (1998).

82. Dressman, D., Yan, H., Traverso, G., Kinzler, K. W. & Vogelstein, B. Transforming single DNA molecules into fluorescent magnetic particles for detection and enumeration of genetic variations. *Proc. Natl. Acad. Sci. U. S. A.* **100,** 8817–8822 (2003).

83. Nakano, M. *et al.* Single-molecule PCR using water-in-oil emulsion. *J. Biotechnol.* **102,** 117–124 (2003).

84. Bremond, N., Thiam, A. R. & Bibette, J. Decompressing emulsion droplets favors

coalescence. *Phys. Rev. Lett.* **100,** 1–4 (2008).

85. Mazutis, L., Baret, J.-C. & Griffiths, A. D. A fast and efficient microfluidic system for highly selective one-to-one droplet fusion. *Lab Chip* **9,** 2665–2672 (2009).

86. Mazutis, L. & Griffiths, A. D. Selective droplet coalescence using microfluidic systems. *Lab Chip* **12,** 1800–1806 (2012).

87. Fidalgo, L. M., Abell, C. & Huck, W. T. S. Surface-induced droplet fusion in microfluidic devices. *Lab Chip* **7,** 984–986 (2007).

88. Chabert, M., Dorfman, K. D. & Viovy, J.-L. L. Droplet fusion by alternating current (AC) field electrocoalescence in microchannels. *Electrophoresis* **26,** 3706–3715 (2005).

89. Baroud, C. N., de Saint Vincent, M. R. & Delville, J.-P. An optical toolbox for total control of droplet microfluidics. *Lab Chip* **7,** 1029–1033 (2007).

90. Link, D. R. *et al.* Electric Control of Droplets in Microfluidic Devices. *Angew. Chemie Int. Ed.* **45,** 2556–2560 (2006).

91. Tewhey, R. *et al.* Microdroplet-based PCR enrichment for large-scale targeted sequencing. *Nat. Biotechnol.* **27,** 1025–1031 (2009).

92. Thiam, A. R., Bremond, N. & Bibette, J. Breaking of an emulsion under an ac electric field. *Phys. Rev. Lett.* **102,** 1–4 (2009).

93. Ahn, K., Agresti, J., Chong, H., Marquez, M. & Weitz, D. A. Electrocoalescence of drops synchronized by size-dependent flow in microfluidic channels. *Appl. Phys. Lett.* **88,** 264105 (2006).

94. Mary, P., Chen, A., Chen, I., Abate, A. R. & Weitz, D. A. On-chip background noise reduction for cell-based assays in droplets. *Lab Chip* **11,** 2066–2070 (2011).

95. Fallah-Araghi, A., Baret, J.-C., Ryckelynck, M. & Griffiths, A. D. A completely in vitro

ultrahigh-throughput droplet-based microfluidic screening system for protein engineering and directed evolution. *Lab Chip* **12,** 882–91 (2012).

96. Tan, W.-H. & Takeuchi, S. Timing controllable electrofusion device for aqueous droplet-based microreactors. *Lab Chip* **6,** 757–63 (2006).

97. Zagnoni, M. & Cooper, J. M. On-chip electrocoalescence of microdroplets as a function of voltage, frequency and droplet size. *Lab Chip* **9,** 2652–2658 (2009).

98. Zagnoni, M., Le Lain, G. & Cooper, J. M. Electrocoalescence mechanisms of microdroplets using localized electric fields in microfluidic channels. *Langmuir* **26,** 14443–14449 (2010).

99. Abate, A. R., Hung, T., Mary, P., Agresti, J. J. & Weitz, D. a. High-throughput injection with microfluidics using picoinjectors. *Proc. Natl. Acad. Sci. U. S. A.* **107,** 19163–19166 (2010).

100. Tice, J. D., Song, H., Lyon, A. D. & Ismagilov, R. F. Formation of Droplets and Mixing in Multiphase Microfluidics at Low Values of the Reynolds and the Capillary Numbers. *Langmuir* **19,** 9127–9133 (2003).

101. Handique, K. & Burns, M. A. Mathematical modeling of drop mixing in a slit-type microchannel. *J. Micromechanics Microengineering* **11,** 548–554 (2001).

102. Bringer, M. R., Gerdts, C. J., Song, H., Tice, J. D. & Ismagilov, R. F. Microfluidic systems for chemical kinetics that rely on chaotic mixing in droplets. *Phil. Trans. R. Soc. Lond. A* **362,** 1087–1104 (2004).

103. Jiang, L., Zeng, Y., Zhou, H., Qu, J. Y. & Yao, S. Visualizing millisecond chaotic mixing dynamics in microdroplets: A direct comparison of experiment and simulation. *Biomicrofluidics* **6,** (2012).

104. Song, H., Tice, J. D. & Ismagilov, R. F. A Microfluidic System for Controlling Reaction Networks in Time - Song - 2003 - Angewandte Chemie - Wiley Online Library. *Angew. Chemie* **115,** 792–796 (2003).

105. Kiss, M. M. *et al.* High-Throughput Quantitative Polymerase Chain Reaction in Picoliter Droplets. *Anal. Chem.* **80,** 8975–8981 (2008).

106. Theberge, A. B. *et al.* Microfluidic platform for combinatorial synthesis in picolitre droplets. *Lab Chip* **12,** 1320–1326 (2012).

107. Leng, X., Zhang, W., Wang, C., Cui, L. & Yang, C. J. Agarose droplet microfluidics for highly parallel and efficient single molecule emulsion PCR. *Lab Chip* **10,** 2841–2843 (2010).

108. Novak, R. *et al.* Single-cell multiplex gene detection and sequencing with microfluidically generated agarose emulsions. *Angew. Chemie - Int. Ed.* **124,** 410–415 (2011).

109. Sepp, A., Tawfik, D. S. & Griffiths, A. D. Microbead display by in vitro compartmentalisation: selection for binding using flow cytometry. *FEBS Lett.* **532,** 455–458 (2002).

110. Levy, M., Griswold, K. E. & Ellington, A. D. Direct selection of trans-acting ligase ribozymes by in vitro compartmentalization. *RNA* **11,** 1555–1562 (2005).

111. Diehl, F. *et al.* BEAMing: single-molecule PCR on microparticles in water-in-oil emulsions. *Nat. Methods* **3,** 551–559 (2006).

112. Margulies, M. *et al.* Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437,** 376–380 (2005).

113. Fidalgo, L. M. *et al.* From Microdroplets to Microfluidics: Selective Emulsion Separation in Microfluidic Devices. *Angew. Chemie Int. Ed.* **47,** 2042–2045 (2008).

114. Pan, X., Zeng, S., Zhang, Q., Lin, B. & Qin, J. Sequential microfluidic droplet processing for rapid DNA extraction. *Electrophoresis* **32,** 3399–3405 (2011).

115. Adamson, D. N., Mustafi, D., Zhang, J. X. J., Zheng, B. & Ismagilov, R. F. Production of arrays of chemically distinct nanolitre plugs via repeated splitting in microfluidic devices. *Lab Chip* **6,** 1178 (2006).

116. Liu, W., Kim, H. J., Lucchetta, E. M., Du, W. & Ismagilov, R. F. Isolation, incubation, and parallel functional testing and identification by FISH of rare microbial single-copy cells from multi-species mixtures using the combination of chemistrode and stochastic confinement. *Lab Chip* **9,** 2153–2162 (2009).

117. Clausell-Tormos, J., Griffiths, A. D. & Merten, C. A. An automated two-phase microfluidic system for kinetic analyses and the screening of compound libraries. *Lab Chip* **10,** 1302–1307 (2010).

118. Sgro, A. E., Allen, P. B. & Chiu, D. T. Thermoelectric manipulation of aqueous droplets in microfluidic devices. *Anal. Chem.* **79,** 4845–4851 (2007).

119. Stan, C. A. *et al.* A microfluidic apparatus for the study of ice nucleation in supercooled water drops. *Lab Chip* **9,** 2293–2305 (2009).

120. Song, H. & Ismagilov, R. F. Millisecond Kinetics on a Microfluidic Chip Using Nanoliters of Reagents. *J. Am. Chem. Soc.* **125,** 14613–14619 (2003).

121. Chanasakulniyom, M. *et al.* Expression of membrane-associated proteins within single emulsion cell facsimiles. *Analyst* **137,** 2939 (2012).

122. Frenz, L., Blank, K., Brouzes, E. & Griffiths, A. D. Reliable microfluidic on-chip incubation of droplets in delay-lines. *Lab Chip* **9,** 1344–8 (2009).

123. Miller, O. J. *et al.* High-resolution dose – response screening using droplet-based

microfluidics. *Proc. Natl. Acad. Sci.* **109,** 378–83 (2012).

124. Hofmann, T. W., Hänselmann, S., Janiesch, J.-W., Rademacher, A. & Böhm, C. H. J. Applying microdroplets as sensors for label-free detection of chemical reactions. *Lab Chip* **12,** 916 (2012).

125. Edd, J. F., Humphry, K. J., Irimia, D., Weitz, D. a & Toner, M. Nucleation and solidification in static arrays of monodisperse drops. *Lab Chip* **9,** 1859–1865 (2009).

126. Hatch, A. C. *et al.* 1-Million droplet array with wide-field fluorescence imaging for digital PCR. *Lab Chip* **11,** 3838–3845 (2011).

127. Huebner, A. *et al.* Static microdroplet arrays: a microfluidic device for droplet trapping, incubation and release for enzymatic and cell-based assays. *Lab Chip* **9,** 692–698 (2009).

128. Laval, P., Lisai, N., Salmon, J.-B. & Joanicot, M. A microfluidic device based on droplet storage for screening solubility diagrams. *Lab Chip* **7,** 829–834 (2007).

129. Um, E., Rha, E., Choi, S.-L., Lee, S.-G. & Park, J.-K. Mesh-integrated microdroplet array for simultaneous merging and storage of single-cell droplets. *Lab Chip* **12,** 1594–1597 (2012).

130. Boukellal, H., Selimović, S., Jia, Y., Cristobal, G. & Fraden, S. Simple, robust storage of drops and fluids in a microfluidic device. *Lab Chip* **9,** 331–8 (2009).

131. Lau, B. T. C., Baitz, C. A., Dong, X. P. & Hansen, C. L. A Complete Microfluidic Screening Platform for Rational Protein Crystallization. *J. Am. Chem. Soc.* **129,** 454–455 (2007).

132. Joensson, H. N. *et al.* Detection and Analysis of Low-Abundance Cell-Surface Biomarkers Using Enzymatic Amplification in Microfluidic Droplets. *Angew. Chemie* **121,** 2556–2559 (2009).

133. Pekin, D. *et al.* Quantitative and sensitive detection of rare mutations using droplet-based microfluidics. *Lab Chip* **11,** 2156–2166 (2011).

134. Mazutis, L. *et al.* Multi-step microfluidic droplet processing: kinetic analysis of an in vitro translated enzyme. *Lab Chip* **9,** 2902–2908 (2009).

135. Hatakeyama, T., Chen, D. L. & Ismagilov, R. F. Microgram-scale testing of reaction conditions in solution using nanoliter plugs in microfluidics with detection by MALDI-MS. *J. Am. Chem. Soc.* **128,** 2518–2519 (2006).

136. Zheng, B. & Ismagilov, R. F. A microfluidic approach for screening submicroliter volumes against multiple reagents by using preformed arrays of nanoliter plugs in a three-phase liquid/liquid/gas flow. *Angew. Chemie - Int. Ed.* **44,** 2520–2523 (2005).

137. Mazutis, L. & Griffiths, A. D. Preparation of monodisperse emulsions by hydrodynamic size fractionation. *Appl. Phys. Lett.* **95,** 204103 (2009).

138. Zhang, K. *et al.* On-chip manipulation of continuous picoliter-volume superparamagnetic droplets using a magnetic force. *Lab Chip* **9,** 2992–2999 (2009).

139. Abate, A. R., Agresti, J. J. & Weitz, D. A. Microfluidic sorting with high-speed single-layer membrane valves. *Appl. Phys. Lett.* **96,** 203509 (2010).

140. Franke, T., Abate, A. R., Weitz, D. A. & Wixforth, A. Surface acoustic wave (SAW) directed droplet flow in microfluidics for PDMS devices. *Lab Chip* **9,** 2625–2627 (2009).

141. Ahn, K. *et al.* Dielectrophoretic manipulation of drops for high-speed microfluidic sorting devices. *Appl. Phys. Lett.* **88,** 24104 (2006).

142. Baret, J. *et al.* Fluorescence-activated droplet sorting (FADS): efficient microfluidic cell sorting based on enzymatic activity. *Lab Chip* **9,** 1850–1858 (2009).

143. Agresti, J. J. *et al.* Ultrahigh-throughput screening in drop-based microfluidics for directed

evolution. *Proc. Natl. Acad. Sci.* **107,** 4004–4009 (2010).

144. Bernath, K. *et al.* In vitro compartmentalization by double emulsions: sorting and gene enrichment by fluorescence activated cell sorting. *Anal. Biochem.* **325,** 151–157 (2004).

145. Zheng, B., Roach, L. S. & Ismagilov, R. F. Screening of protein crystallization conditions on a microfluidic chip using nanoliter-size droplets. *J. Am. Chem. Soc.* **125,** 11170–11171 (2003).

146. Kintses, B., van Vliet, L. D., Devenish, S. R. A. & Hollfelder, F. Microfluidic droplets: New integrated workflows for biological experiments. *Curr. Opin. Chem. Biol.* **14,** 548–555 (2010).

147. Pompano, R. R., Liu, W., Du, W. & Ismagilov, R. F. Microfluidics Using Spatially Defined Arrays of Droplets in One, Two, and Three Dimensions. *Annu. Rev. Anal. Chem.* **4,** 59–81 (2011).

148. Chen, D. L. & Ismagilov, R. F. Microfluidic cartridges preloaded with nanoliter plugs of reagents: an alternative to 96-well plates for screening. *Curr. Opin. Chem. Biol.* **10,** 226–231 (2006).

149. Schonbrun, E., Abate, A. R., Steinvurzel, P. E., Weitz, D. A. & Crozier, K. B. High-throughput fluorescence detection using an integrated zone-plate array. *Lab Chip* **10,** 852–856 (2010).

150. Srisa-Art, M., DeMello, A. J. & Edel, J. B. High-throughput DNA droplet assays using picoliter reactor volumes. *Anal. Chem.* **79,** 6682–6689 (2007).

151. Zheng, B., Tice, J. D. & Ismagilov, R. F. Formation of Droplets of Alternating Composition in Microfluidic Channels and Applications to Indexing of Concentrations in Droplet-Based Assays. *Anal. Chem.* **76,** 4977–4982 (2004).

152. Baret, J.-C. C., Beck, Y., Billas-Massobrio, I., Moras, D. & Griffiths, A. D. Quantitative cell-based reporter gene assays using droplet-based microfluidics. *Chem. Biol.* **17,** 528–536 (2010).

153. Melkko, S., Zhang, Y., Dumelin, C. E., Scheuermann, J. & Neri, D. Isolation of High-Affinity Trypsin Inhibitors from a DNA-Encoded Chemical Library. *Angew. Chemie Int. Ed.* **46,** 4671–4674 (2007).

154. Portney, N. G., Wu, Y., Quezada, L. K., Lonardi, S. & Ozkan, M. Length-Based Encoding of Binary Data in DNA. *Langmuir* **24,** 1613–1616 (2008).

155. Lorenceau, E. *et al.* Generation of Polymerosomes from Double-Emulsions. *Langmuir* **21,** 9183–9186 (2005).

156. Shah, R. K., Kim, J.-W., Agresti, J. J., Weitz, D. A. & Chu, L.-Y. Fabrication of monodisperse thermosensitive microgels and gel capsules in microfluidic devices. *Soft Matter* **4,** 2303–2309 (2008).

157. Ji, X.-H. *et al.* Integrated parallel microfluidic device for simultaneous preparation of multiplex optical-encoded microbeads with distinct quantum dot barcodes. *J. Mater. Chem.* **21,** 13380 (2011).

158. Theilacker, N., Roller, E. E., Barbee, K. D., Franzreb, M. & Huang, X. Multiplexed protein analysis using encoded antibody-conjugated microbeads. *J. R. Soc. Interface* **8,** 1104–1113 (2011).

159. Zhang, P., He, Y., Ruan, Z., Chen, F. F. & Yang, J. Fabrication of quantum dots-encoded microbeads with a simple capillary fluidic device and their application for biomolecule detection. *J. Colloid Interface Sci.* **385,** 8–14 (2012).

160. Miller, M. B. & Tang, Y.-W. Basic Concepts of Microarrays and Potential Applications in

Clinical Microbiology. *Clin. Microbiol. Rev.* **22,** 611–633 (2009).

161. Niu, X. Z. *et al.* Droplet-based compartmentalization of chemically separated components in two-dimensional separations. *Chem. Commun. (Camb).* 6159–61 (2009). doi:10.1039/b918100h

162. Ji, J. *et al.* Proteolysis in microfluidic droplets: an approach to interface protein separation and peptide mass spectrometry. *Lab Chip* **12,** 2625 (2012).

163. Liu, S. *et al.* The electrochemical detection of droplets in microfluidic devices. *Lab Chip* **8,** 1937–1942 (2008).

164. Luo, C. *et al.* Picoliter-volume aqueous droplets in oil: Electrochemical detection and yeast cell electroporation. *Electrophoresis* **27,** 1977–1983 (2006).

165. Cristobal, G. *et al.* On-line laser Raman spectroscopic probing of droplets engineered in microfluidic devices. *Lab Chip* **6,** 1140–1146 (2006).

166. Strehle, K. R. *et al.* A reproducible surface-enhanced Raman spectroscopy approach. Online SERS measurements in a segmented microfluidic system. *Anal. Chem.* **79,** 1542–1547 (2007).

167. Wang, G. *et al.* Surface-enhanced Raman scattering in nanoliter droplets: towards high-sensitivity detection of mercury (II) ions. *Anal. Bioanal. Chem.* **394,** 1827–1832 (2009).

168. Ackermann, K. R., Henkel, T. & Popp, J. Quantitative online detection of low-concentrated drugs via a SERS microfluidic system. *ChemPhysChem* **8,** 2665–2670 (2007).

169. März, A. *et al.* Towards a quantitative SERS approach - online monitoring of analytes in a microfluidic system with isotope-edited internal standards. *J. Biophotonics* **2,** 232–242 (2009).

170. Cecchini, M. P. *et al.* Ultrafast surface enhanced resonance raman scattering detection in droplet-based microfluidic systems. *Anal. Chem.* **83,** 3076–3081 (2011).

171. Choi, J.-W., Kang, D.-K., Park, H., DeMello, A. J. & Chang, S.-I. High-throughput analysis of protein-protein interactions in picoliter-volume droplets using fluorescence polarization. *Anal. Chem.* **84,** 3849–54 (2012).

172. Nguyen, N. T., Lassemono, S. & Chollet, F. A. Optical detection for droplet size control in microfluidic droplet-based analysis systems. *Sensors Actuators, B Chem.* **117,** 431–436 (2006).

173. Srisa-Art, M., Dyson, E. C., DeMello, A. J. & Edel, J. B. Monitoring of real-time streptavidin-biotin binding kinetics using droplet microfluidics. *Anal. Chem.* **80,** 7063–7067 (2008).

174. Srisa-Art, M., Demello, A. J. & Edel, J. B. Fluorescence lifetime imaging of mixing dynamics in continuous-flow microdroplet reactors. *Phys. Rev. Lett.* **101,** 1–4 (2008).

175. Srisa-Art, M., deMello, A. J. & Edel, J. B. High-throughput confinement and detection of single DNA molecules in aqueous microdroplets. *Chem. Commun. (Camb).* **44,** 6548–6550 (2009).

176. Niu, X., Gielen, F., Edel, J. B. & deMello, A. J. A microdroplet dilutor for high-throughput screening. *Nat. Chem.* **3,** 437–442 (2011).

177. Srisa-Art, M. *et al.* Identification of rare progenitor cells from human periosteal tissue using droplet microfluidics. *Analyst* **134,** 2239–45 (2009).

178. Granieri, L., Baret, J. C., Griffiths, A. D. & Merten, C. A. High-Throughput Screening of Enzymes by Retroviral Display Using Droplet-Based Microfluidics. *Chem. Biol.* **17,** 229–235 (2010).

179. Wang, J. *et al.* Fast self-assembly kinetics of quantum dots and a dendrimeric peptide ligand. *Langmuir* **28,** 7962–7966 (2012).

180. Srisa-Art, M. *et al.* Analysis of Protein-Protein Interactions by Using Droplet-Based Microfluidics. *ChemBioChem* **10,** 1605–1611 (2009).

181. Casadevall i Solvas, X., Srisa-Art, M., deMello, A. J. & Edel, J. B. Mapping of fluidic mixing in microdroplets with 1 micros time resolution using fluorescence lifetime imaging. *Anal. Chem.* **82,** 3950–6 (2010).

182. Cherry, J. R. & Fidantsef, A. L. Directed evolution of industrial enzymes: An update. *Curr. Opin. Biotechnol.* **14,** 438–443 (2003).

183. Dove, A. Drug screening - beyond the bottleneck. *Nat. Biotechnol.* **17,** 859–863 (1999).

184. Gong, Z. *et al.* Drug effects analysis on cells using a high throughput microfluidic chip. *Biomed. Microdevices* **13,** 215–219 (2011).

185. Zhan, Y., Wang, J., Bao, N. & Lu, C. Electroporation of cells in microfluidic droplets. *Anal. Chem.* **81,** 2027–2031 (2009).

186. Rotman, B. Measurement of activity of single molecules of beta-D-galactosidase. *Proc. Natl. Acad. Sci. U. S. A.* **47,** 1981–1991 (1961).

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

188. Zhang, H., Jenkins, G., Zou, Y., Zhu, Z. & Yang, C. J. Massively parallel single-molecule and single-cell emulsion reverse transcription polymerase chain reaction using agarose droplet microfluidics. *Anal. Chem.* **84,** 3599–3606 (2012).

189. Boedicker, J. Q., Li, L., Kline, T. R. & Ismagilov, R. F. Detecting bacteria and determining their susceptibility to antibiotics by stochastic confinement in nanoliter

droplets using plug-based microfluidics. *Lab Chip* **8,** 1265–1272 (2008).

190. Chao, W.-C. *et al.* Control of concentration and volume gradients in microfluidic droplet arrays for protein crystallization screening. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* **4,** 2623–6 (2004).

191. Maeki, M. *et al.* X-ray diffraction of protein crystal grown in a nano-liter scale droplet in a microchannel and evaluation of its applicability. *Anal. Sci.* **28,** 65 (2012).

192. Chayen, N. E. & Saridakis, E. Protein crystallization: from purified protein to diffraction-quality crystal. *Nat. Methods* **5,** 147–153 (2008).

193. Zhang, W. Y. *et al.* Highly parallel single-molecule amplification approach based on agarose droplet polymerase chain reaction for efficient and cost-effective aptamer selection. *Anal. Chem.* **84,** 350–355 (2012).

194. Konry, T., Dominguez-Villar, M., Baecher-Allan, C., Hafler, D. A. & Yarmush, M. L. Droplet-based microfluidic platforms for single T cell secretion analysis of IL-10 cytokine. *Biosens. Bioelectron.* **26,** 2707–2710 (2011).

195. Zhong, Q. *et al.* Multiplex digital PCR: breaking the one target per color barrier of quantitative PCR. *Lab Chip* **11,** 2167–2174 (2011).

196. Mazutis, L. *et al.* Droplet-based microfluidic systems for high-throughput single DNA molecule isothermal amplification and analysis. *Anal. Chem.* **81,** 4813–4821 (2009).

197. Whale, A. S. *et al.* Comparison of microfluidic digital PCR and conventional quantitative PCR for measuring copy number variation. *Nucleic Acids Res.* **40,** (2012).

198. Heyries, K. a *et al.* Megapixel digital PCR. *Nat. Methods* **8,** 649–651 (2011).

199. Hussein, S. M. *et al.* Copy number variation and selection during reprogramming to pluripotency. *Nature* **471,** 58–62 (2011).

200. Zhu, Z. *et al.* Single-molecule emulsion PCR in microfluidic droplets. *Anal. Bioanal. Chem.* **403,** 2127–2143 (2012).

201. Xia, Y. & Whitesides, G. M. Soft Lithography. *Annu. Rev. Mater. Sci.* **28,** 153–184 (1998).

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

203. Duffy, D. C., Schueller, O. J. a, Brittain, S. T. & Whitesides, G. M. Rapid prototyping of microfluidic switches in poly(dimethyl siloxane) and their actuation by electro-osmotic flow. *J. Micromechanics Microengineering* **9,** 211–217 (1999).

204. Samel, B., Chowdhury, M. K. & Stemme, G. The fabrication of microfluidic structures by means of full-wafer adhesive bonding using a poly(dimethylsiloxane) catalyst. *J. Micromechanics Microengineering* **17,** 1710–1714 (2007).

205. McEwan, A. D. The peeling of a flexible strip attached by a cavitating viscous adhesive. *Rheol. Acta* **5,** 205–211 (1966).

206. Saffman, P. G. & Taylor, G. I. The penetration of a fluid into a porous medium or Hele-Shaw cell containing a more viscous liquid. *Proc. R. Soc. A* **245,** 312–329 (1958).

207. Eddings, M. a, Johnson, M. a & Gale, B. K. Determining the optimal PDMS–PDMS bonding technique for microfluidic devices. *J. Micromechanics Microengineering* **18,** 67001 (2008).

208. Bhattacharya, S., Datta, A., Berg, J. M. & Gangopadhyay, S. Studies on surface wettability of poly(dimethyl) siloxane (PDMS) and glass under oxygen-plasma treatment and correlation with bond strength. *J. Microelectromechanical Syst.* **14,** 590–597 (2005).

209. Edgar, J. S. *et al.* Capillary electrophoresis separation in the presence of an immiscible

boundary for droplet analysis. *Anal. Chem.* **78,** 6948–6954 (2006).

210. Abate, A. R. *et al.* Photoreactive coating for high-contrast spatial patterning of microfluidic device wettability. *Lab Chip* **8,** 2157–2160 (2008).

211. Roberts, C. C. *et al.* Comparison of monodisperse droplet generation in flow-focusing devices with hydrophilic and hydrophobic surfaces. *Lab Chip* **12,** 1540 (2012).

212. Bauer, W.-A. C., Fischlechner, M., Abell, C. & Huck, W. T. S. Hydrophilic PDMS microchannels for high-throughput formation of oil-in-water microdroplets and water-in-oil-in-water double emulsions. *Lab Chip* **10,** 1814–9 (2010).

213. Eastburn, D. J. *et al.* Microfluidic droplet enrichment for targeted sequencing. *Nucleic Acids Res.* **43,** e86 (2015).

214. Santelli, C. M. *et al.* Abundance and diversity of microbial life in ocean crust. *Nature* **453,** 653–656 (2008).

215. Metzker, M. L. Sequencing technologies - the next generation. *Nat. Rev. Genet.* **11,** 31–46 (2010).

216. Shendure, J. & Ji, H. Next-generation DNA sequencing. *Nat. Biotechnol.* **26,** 1135–1145 (2008).

217. Loman, N. J. *et al.* High-throughput bacterial genome sequencing: an embarrassment of choice, a world of opportunity. *Nat. Rev. Microbiol.* **10,** 599–606 (2012).

218. Chen, X.-H. *et al.* Structural and functional characterization of three polyketide synthase gene clusters in Bacillus amyloliquefaciens FZB 42. *J. Bacteriol.* **188,** 4024–36 (2006).

219. Arnold, W., Rump, A., Klipp, W., Priefer, U. B. & Pühler, A. Nucleotide sequence of a 24,206-base-pair DNA fragment carrying the entire nitrogen fixation gene cluster of Klebsiella pneumoniae. *J. Mol. Biol.* **203,** 715–738 (1988).

220. Osbourn, A. Secondary metabolic gene clusters: Evolutionary toolkits for chemical innovation. *Trends Genet.* **26,** 449–457 (2010).

221. Fischbach, M. A., Walsh, C. T. & Jon, C. The evolution of gene collectives: How natural selection drives chemical innovation. *Proc. Natl. Acad. Sci.* **105,** 4601–4608 (2008).

222. Li, R. *et al.* De novo assembly of human genomes with massively parallel short read sequencing. *Genome Res.* **20,** 265–272 (2010).

223. Mardis, E. R. The impact of next-generation sequencing technology on genetics. *Trends Genet.* **24,** 133–141 (2008).

224. Mamanova, L. *et al.* Target-enrichment strategies for next- generation sequencing. *Nat. Methods* **7,** 111–118 (2010).

225. Chan, M. *et al.* Development of a next-generation sequencing method for brca mutation screening: A comparison between a high-throughput and a benchtop platform. *J. Mol. Diagnostics* **14,** 602–612 (2012).

226. Tasse, L. *et al.* Functional metagenomics to mine the human gut microbiome for dietary fiber catabolic enzymes Functional metagenomics to mine the human gut microbiome for dietary fiber catabolic enzymes. *Genome Res.* 1605–1612 (2010). doi:10.1101/gr.108332.110

227. Acinas, S. G., Sarma-Rupavtarm, R., Klepac-Ceraj, V. & Polz, M. F. PCR-induced sequence artifacts and bias: Insights from comparison of two 16s rRNA clone libraries constructed from the same sample. *Appl. Environ. Microbiol.* **71,** 8966–8969 (2005).

228. Schloss, P. D., Gevers, D. & Westcott, S. L. Reducing the effects of PCR amplification and sequencing Artifacts on 16s rRNA-based studies. *PLoS One* **6,** (2011).

229. Elnifro, E. M., Ashshi, A. M., Cooper, R. J. & Klapper, P. E. Multiplex PCR :

Optimization and Application in Diagnostic Virology Multiplex PCR : Optimization and Application in Diagnostic Virology. *Clin. Microbiol. Rev.* **13,** 559–570 (2000).

230. Ottesen, E. A., Hong, J. W., Quake, S. & Leadbetter, J. R. Microfluidic digital PCR enables multigene analysis of individual environmental bacteria. *Science (80-.).* 1464–1467 (2006). doi:10.1126/science.1131370

231. Mertes, F. *et al.* Targeted enrichment of genomic DNA regions for next-generation sequencing. *Brief. Funct. Genomics* **10,** 374–386 (2011).

232. Taly, V. *et al.* Multiplex picodroplet digital PCR to detect KRAS mutations in circulating DNA from the plasma of colorectal cancer patients. *Clin. Chem.* **59,** 1722–1731 (2013).

233. Bodi, K. *et al.* Comparison of commercially available target enrichment methods for next-generation sequencing. *J. Biomol. Tech.* **24,** 73–86 (2013).

234. Kim, S. C., Sukovich, D. J. & Abate, A. R. Patterning microfluidic device wettability with spatially-controlled plasma oxidation. *Lab Chip* **15,** 3163–3169 (2015).

235. Lim, S. W. & Abate, A. R. Ultrahigh-throughput sorting of microfluidic drops with flow cytometry. *Lab Chip* **13,** 4563–72 (2013).

236. Mashaghi, S., Abbaspourrad, A., Weitz, D. A. & van Oijen, A. M. Droplet microfluidics: A tool for biology, chemistry and nanotechnology. *TrAC - Trends Anal. Chem.* **82,** 118–125 (2016).

237. Schneeberger, C., Speiser, P., Kury, F. & Zeillinger, R. Quantitative Detection of Reverse Transcriptase-PCR Products by Means of a Novel and Sensitive DNA Stain. *Genome Res.* **4,** 234–238 (1995).

238. Holland, P. M., Abramson, R. D., Watson, R. & Gelfand, D. H. Detection of specific polymerase chain reaction product by utilizing the 5’---3’ exonuclease activity of Thermus

aquaticus DNA polymerase. *Proc Natl Acad Sci U S A* **88,** 7276–7280 (1991).

239. Casamayor, E. O. *et al.* Changes in archaeal, bacterial and eukaryal assemblages along a salinity gradient by comparison of genetic fingerprinting methods in a multipond solar saltern. *Environ. Microbiol.* **4,** 338–348 (2002).

240. Meays, C. L., Broersma, K., Nordin, R. & Mazumder, A. Source tracking fecal bacteria in water: A critical review of current methods. *J. Environ. Manage.* **73,** 71–79 (2004).

241. Goldgar, D. E. *et al.* Integrated evaluation of DNA sequence variants of unknown clinical significance: application to BRCA1 and BRCA2. *Am. J. Hum. Genet.* **75,** 535–544 (2004).

242. Basler, C. F. *et al.* Sequence of the 1918 pandemic influenza virus nonstructural gene (NS) segment and characterization of recombinant viruses bearing the 1918 NS genes. *Proc. Natl. Acad. Sci.* **98,** 2746–2751 (2001).

243. Smith, D. J. *et al.* Mapping the Antigenic and Genetic Evolution of Influenza Virus. *Science (80-.).* **305,** 371–376 (2004).

244. Walther, W. & Stein, U. Viral vectors for gene transfer: a review of their use in the treatment of human diseases. *Drugs* **60,** 249–71 (2000).

245. Kreppel, F. & Kochanek, S. Modification of adenovirus gene transfer vectors with synthetic polymers: a scientific review and technical guide. *Mol. Ther.* **16,** 16–29 (2008).

246. Bohannan, B. J. M. & Lenski, R. E. Linking genetic change to community evolution: Insights from studies of bacteria and bacteriophage. *Ecology Letters* **3,** 362–377 (2000).

247. Suttle, C. a. Viruses in the sea. *Nature* **437,** 356–361 (2005).

248. Carroll, M. W. *et al.* Temporal and spatial analysis of the 2014-2015 Ebola virus outbreak in West Africa. *Nature* **524,** 97–101 (2015).

249. Treanor, J. J. in *Viral Infections of Humans: Epidemiology and Control* (eds. Kaslow, R.

A., Stanberry, L. R. & Le Duc, J. W.) 455–478 (Springer, 2014). doi:10.1007/978-1-4899-7448-8

250. Young, L. S. & Rickinson, A. B. Epstein-Barr virus: 40 years on. *Nat.Rev.Cancer* **4,** 757–768 (2004).

251. Trono, D. *et al.* HIV Persistence and the Prospect of. 174–180 (2010).

252. Peterson, L. R., Jamieson, D. J., Powers, A. M. & Honein, M. A. Zika Virus. *N. Engl. J. Med.* **374,** 1552–1563 (2016).

253. Beerenwinkel, N., Gunthard, H. F., Roth, V. & Metzner, K. J. Challenges and opportunities in estimating viral genetic diversity from next-generation sequencing data. *Front. Microbiol.* **3,** 1–16 (2012).

254. Mokili, J. L., Rohwer, F. & Dutilh, B. E. Metagenomics and future perspectives in virus discovery. *Curr. Opin. Virol.* **2,** 63–77 (2012).

255. Fuhrman, J. a. & Campbell, L. Marine ecology: Microbial microdiversity. *Nature* **393,** 410–411 (1998).

256. Hugenholtz, P. Exploring prokaryotic diversity in the genomic era. *Genome Biol.* **3,** REVIEWS0003 (2002).

257. Breitbart, M. *et al.* Genomic analysis of uncultured marine viral communities. *Proc. Natl. Acad. Sci. U. S. A.* **99,** 14250–5 (2002).

258. Angly, F. E. *et al.* The marine viromes of four oceanic regions. *PLoS Biol.* **4,** 2121–2131 (2006).

259. Edwards, R. A. & Rohwer, F. Viral metagenomics. *Nat. Rev. Microbiol.* **17,** 504–510 (2005).

260. Alkan, C., Sajjadian, S. & Eichler, E. E. Limitations of next-generation genome sequence

assembly. *Nat. Methods* **8,** 61–65 (2010).

261. Kunin, V., Copeland, A., Lapidus, A., Mavromatis, K. & Hugenholtz, P. A bioinformatician's guide to metagenomics. *Microbiol. Mol. Biol. Rev.* **72,** 557–78, Table of Contents (2008).

262. Pappas, D. & Wang, K. Cellular separations: A review of new challenges in analytical chemistry. *Anal. Chim. Acta* **601,** 26–35 (2007).

263. Sidore, A. M., Lan, F., Lim, S. W. & Abate, A. R. Enhanced sequencing coverage with digital droplet multiple displacement amplification. *Nucleic Acids Res.* gkv1493 (2015). doi:10.1093/nar/gkv1493

264. Tran, T. M. *et al.* A nested real-time PCR assay for the quantification of Plasmodium falciparum DNA extracted from dried blood spots. *Malar. J.* **13,** (2014).

265. Thompson, L. R. *et al.* PNAS Plus: Phage auxiliary metabolic genes and the redirection of cyanobacterial host carbon metabolism. *Proceedings of the National Academy of Sciences* **108,** E757–E764 (2011).

266. Hurwitz, B. L., Hallam, S. J. & Sullivan, M. B. Metabolic reprogramming by viruses in the sunlit and dark ocean. *Genome Biol.* **14,** R123 (2013).

267. Rappé, M. S. & Giovannoni, S. J. The uncultured microbial majority. *Annu. Rev. Microbiol.* **57,** 369–394 (2003).

268. Vartoukian, S. R., Palmer, R. M. & Wade, W. G. Strategies for culture of 'unculturable' bacteria. *FEMS Microbiol. Lett.* **309,** 1–7 (2010).

269. Reyes, A., Semenkovich, N. P., Whiteson, K., Rohwer, F. & Gordon, J. I. Going viral: next-generation sequencing applied to phage populations in the human gut. *Nat. Rev. Microbiol.* **10,** 607–617 (2012).

270. Dhillon, V. & Li, X. Single-Cell Genome Sequencing for Viral-Host Interactions. *J. Comput. Sci. Syst. Biol.* **8,** 160–165 (2015).

271. Roux, S. *et al.* Ecology and evolution of viruses infecting uncultivated SUP05 bacteria as revealed by single-cell and environmental genomics. *Elife* **in review,** 1–20 (2014).

272. Labonté, J. M. *et al.* Single-cell genomics-based analysis of virus–host interactions in marine surface bacterioplankton. *ISME J.* 1–14 (2015). doi:10.1038/ismej.2015.48

273. Andersson, A. F. & Banfield, J. F. Virus population dynamics and acquired virus resistance in natural microbial communities. *Science* **320,** 1047–1050 (2008).

274. Weinberger, A. D. *et al.* Persisting viral sequences shape microbial CRISPR-based immunity. *PLoS Comput. Biol.* **8,** (2012).

275. Allers, E. *et al.* Single-cell and population level viral infection dynamics revealed by phageFISH, a method to visualize intracellular and free viruses. *Environ. Microbiol.* **15,** 2306–2318 (2013).

276. Deng, L. *et al.* Contrasting life strategies of viruses that infect photo- and heterotrophic bacteria, as revealed by viral tagging. *MBio* **3,** 1–8 (2012).

277. Tadmor, A. D., Ottesen, E. A., Leadbetter, J. R. & Phillips, R. Probing individual environmental bacteria for viruses by using microfluidic digital PCR. *Science* **333,** 58–62 (2011).

278. Dang, V. T. & Sullivan, M. B. Emerging methods to study bacteriophage infection at the single-cell level. *Front. Microbiol.* **5,** 1–8 (2014).

279. Lim, S. W., Tran, T. M. & Abate, A. R. PCR-Activated Cell Sorting for Cultivation-Free Enrichment and Sequencing of Rare Microbes. *PLoS One* **10,** e0113549 (2015).

280. Eastburn, D. J., Sciambi, A. & Abate, A. R. Identification and genetic analysis of cancer

cells with PCR-activated cell sorting. *Nucleic Acids Res.* **42,** 1–10 (2014).

281. Karam, J. D. *Molecular Biology of Bacteriophage T4*. (ASM Press, 1994).

282. Arber, W., Enquist, L., Hohn, B., Murray, N. & Murray, K. *Experimental methods for use with lambda*. *Lambda II* **1,** (1983).

283. Hafenstein, S. & Fane, B. a. phi X174 genome-capsid interactions influence the biophysical properties of the virion: evidence for a scaffolding-like function for the genome during the final stages of morphogenesis. *J. Virol.* **76,** 5350–5356 (2002).

284. Mazutis, L. *et al.* Single-cell analysis and sorting using droplet-based microfluidics. *Nat. Protoc.* **8,** 870–91 (2013).

285. Sciambi, A. & Abate, A. R. Accurate microfluidic sorting of droplets at 30 kHz. *Lab Chip* **15,** 47–51 (2015).

286. Sciambi, A. & Abate, A. R. Generating electric fields in PDMS microfluidic devices with salt water electrodes. *Lab Chip* **14,** 2605–2609 (2014).

287. Tran, T. M., Lan, F., Thompson, C. S. & Abate, a R. From tubes to drops: droplet-based microfluidics for ultrahigh-throughput biology. *J. Phys. D. Appl. Phys.* **46,** 114004 (2013).

288. Turner, E. H., Ng, S. B., Nickerson, D. a & Shendure, J. Methods for genomic partitioning. *Annu. Rev. Genomics Hum. Genet.* **10,** 263–284 (2009).

289. Diemer, G. S. & Stedman, K. M. A novel virus genome discovered in an extreme environment suggests recombination between unrelated groups of RNA and DNA viruses. *Biol. Direct* **7,** 13 (2012).

290. Stedman, K. Mechanisms for RNA capture by ssDNA viruses: Grand theft RNA. *J. Mol. Evol.* **76,** 359–364 (2013).

291. Archin, N. M., Sung, J. M., Garrido, C., Soriano-Sarabia, N. & Margolis, D. M.

Eradicating HIV-1 infection: seeking to clear a persistent pathogen. *Nat. Rev. Microbiol.* **12,** 750–764 (2014).

292. Pasternak, A. O. *et al.* Highly sensitive methods based on seminested real-time reverse transcription-PCR for quantitation of human immunodeficiency virus type 1 unspliced and multiply spliced RNA and proviral DNA. *J. Clin. Microbiol.* **46,** 2206–2211 (2008).

293. Deeks, S. G. *et al.* Towards an HIV cure: a global scientific strategy. *Nat. Rev. Immunol.* **12,** 607–614 (2012).

294. Eriksson, S. *et al.* Comparative Analysis of Measures of Viral Reservoirs in HIV-1 Eradication Studies. *PLoS Pathog.* **9,** (2013).

295. Shan, L. *et al.* A novel PCR assay for quantification of HIV-1 RNA. *J. Virol.* **87,** 6521–5 (2013).

296. Fischer, M. *et al.* Biphasic decay kinetics suggest progressive slowing in turnover of latently HIV-1 infected cells during antiretroviral therapy. *Retrovirology* **5,** 107 (2008).

297. Gallastegui, E. *et al.* Combination of Biological Screening in a Cellular Model of Viral Latency and Virtual Screening Identifies Novel Compounds That Reactivate HIV-1. *J. Virol.* **86,** 3795–3808 (2012).

298. Fischer, M. *et al.* Cellular viral rebound after cessation of potent antiretroviral therapy predicted by levels of multiply spliced HIV-1 RNA encoding nef. *J. Infect. Dis.* **190,** 1979–1988 (2004).

299. Clavel, F. *et al.* Genetic recombination of human immunodeficiency virus. *J. Virol.* **63,** 1455–1459 (1989).

300. Siliciano, J. D. & Siliciano, R. F. Enhanced Culture Assay for Detection and Quantitation of Latently Infected , Replication-Competent Virus in HIV-1-Infected Individuals. *Mol.*

Biol. **304,** 3–15

301. Ho, Y. C. *et al.* XReplication-competent noninduced proviruses in the latent reservoir increase barrier to HIV-1 cure. *Cell* **155,** 540–551 (2013).

302. Laird, G. M. *et al.* Rapid Quantification of the Latent Reservoir for HIV-1 Using a Viral Outgrowth Assay. *PLoS Pathog.* **9,** (2013).

Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.

Author Signature___  Date December 8, 2016