Title
Microfluidic tools for connecting single-cell optical and gene expression phenotype

Author
Zhang, Jesse Qiuxu

Publication Date
2020

Peer reviewed|Thesis/dissertation

Microfluidic tools for connecting single-cell optical and gene expression phenotype

by

Jesse Qiuxu Zhang

DISSERTATION

Submitted in partial satisfaction of the requirements for degree of

DOCTOR OF PHILOSOPHY

in

Bioengineering

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO
AND
UNIVERSITY OF CALIFORNIA, BERKELEY

Approved:

DocuSigned by:
Adam R. Abate
FC0EE10723AF40D...

Adam R. Abate
Chair

DocuSigned by:

Zev Gartner

DocuSigned by:
...4DD...

Aaron Streets

DocuSigned by:
...47B...

Chun Jimmie Ye
A758AEFF054747C...

Committee Members

Acknowledgements

While this dissertation represents an individual contribution to Science, it would not have been possible without the support of many individuaSls along the way.

I have been fortunate to have been co-mentored during graduate school by Prof. Adam Abate and Prof. Zev Gartner. I thank Adam for teaching me how to be a more effective scientist and communicator, and being fully invested in my training since when I first entered his lab four years ago. I thank Zev for being always open to new ideas, always listening and providing great insight with regards to my science. He has contributed directly or indirectly to every one of my successes during my time at UCSF.

I thank my qualifying examination committee, which consisted of Prof. Joe DeRisi, Prof. Aaron Streets, Prof. Jimmie Ye, and was chaired by Prof. Bo Huang. They provided crucial guidance during the early stages of my dissertation work. I am thankful for Aaron and Jimmie sticking around to be part of my dissertation committee. I had many great discussions with Aaron about our research and the overall direction of the microfluidics and single-cell genomics field. In my discussions with Jimmie, he has always had a keen eye to how the technologies that I was developing could be even more impactful. I thank the two of them for always believing in my potential and helping guide me through the dissertation process and beyond.

There are many members, past and current, of the Abate lab I would like to thank for making my time spent there productive and enjoyable. I thank Russell Cole and Christian Siltanen for mentoring me and helping me sow the seeds for what would become my first completed project. I thank Shiyang Tang for teaching me many of the nuances of photolithography and microfluidic device fabrication. I thank Leqian Liu for teaching me the basics of molecular biology and always being willing to lend a hand with my experiments. I have

had the pleasure of working closely with Kai-Chun Chang over the last years of my dissertation. I thank him for being patient with me as I brought him up to speed in the lab, being a great contributor to my later successes in the lab, and always being willing to talk about science or anything else in life. Overall, because of these people and everyone else in lab, I have had the pleasure of working in such an engaging learning environment with a helpful and collaborative culture.

One advantage of being co-mentored is that I have been able to contribute, learn, and grow from two labs. The Gartner lab has been a great second lab at UCSF, with a friendly and helpful culture filled with great people. I would, in particular, like to thank several people in lab. I thank Alex Hughes for mentoring me back when I was rotating with the lab and providing great discussion about the process of being a scientist, back when I was just starting grad school. I thank Katie Cabral for being a great person to talk with and learn a different perspective about microfluidics. Alex and Katie have motivated me to push the boundaries of microfluidics with my science. I thank Chris McGinnis for all the great technology development discussions he had and always being willing to answer my bioinformatic questions. Lastly, my time in the Gartner lab has shown me that science and fun can co-exist both inside and outside of work.

My journey to completing this dissertation started much before I arrived at UCSF. I would like to thank some mentors from earlier in my scientific career. Back when I was in high school, I had the pleasure of working with Prof. Naoyuki Miura and his team at the Hamamatsu University School of Medicine. It was my first research experience, and gave me this passion and excitement for performing life sciences research that I brought with me to UCSF. I thank Prof. Christian Kastrup at the University of British Columbia for taking me up for a summer research position in 2012. He and James Baylis, who mentored me during the position, gave me

the tools and freedom for me to grow as a scientist. During my undergraduate time at Northwestern University, I had the pleasure of working with Prof. Lonnie Shea and his team. I like to thank him and all the members of his lab for making my time there so productive and enjoyable. I thank Michael Gower, now a professor at the University of South Carolina, for being a great mentor, showing me how to perform rigorous science, and most importantly showing me that one could keep up a work life balance while doing so. I thank Jeffrey Liu for being a collegial project partner. The positive experiences I had working with Michael and Jeffrey were a large factor to motivating me to pursue a PhD immediately after college. I thank Brian Aguado, soon to be professor at UCSD, for mentoring me during the grad school application process during my last year of undergrad. His honest and thoughtful assistance with my application package I believe are a large part of why I had the opportunity to visit and attend UCSF.

Behind every success of mine during grad school were tens of failures. Having people outside of science to support me through those failures, and help me forget and most past them was a huge contributor to why I made it far enough in grad school to compile this dissertation. I thank Alex, Aram, Brian, Kimberly, Jonathan, Riya, and all the other friends I have in the Bay Area for helping me explore all that the area has to offer. I thank Andrew, Jeremi, Jonathan, and Justin for being there with me since our first days of college and the now five years after graduation. I thank Kevin for being a great friend with whom to navigate our surprisingly parallel post-graduation trajectories in life. For the past two years, I have also had amazing support from my partner Jocelyn, who has seen and been there to support me ever since I passed my qualifying exam.

I lastly want to thank my father Xu Zhang for always encouraging me to be adventurous and to forge my own path in life and my mother Hui Zhu for her unconditional support and love through the ups and downs of the past five years of graduate school and the 21 years prior to that.

Contributions

Elements of this dissertation have been published elsewhere or are in submission for publication at peer reviewed journals. Chapters 2 and 3 have been adapted from published articles in peer reviewed journals. Chapter 4 is adapted from a manuscript in submission at a peer reviewed journal.

Microfluidic tools for connecting single-cell optical and gene expression phenotype

Jesse Qiuxu Zhang

Abstract

The single cell is the fundamental unit of biology. Understanding how the identity of individual cells in multicellular organisms contribute to their function remains a key question in Biology. Traditionally, most observations of cells were made through imaging-based techniques. Today, advances in Next Generation Sequencing have led to the widespread adoption of sequencing-based techniques for investigating the genotype and phenotype at single-cell resolution. Microfluidics, including droplet-based microfluidics, have been instrumental in the most successful commercial single-cell genomics platforms.

Integrating sequencing and imaging techniques will provide additional information than either of the techniques alone. Both single-cell imaging and genomics techniques measure orthogonal targets, and when combined reveal additional insights into cellular function. However, when performing sequential single-cell assays, there currently exists a tradeoff between throughput and information content. This dissertation will describe progress made towards reducing that gap. I will describe novel microfluidic platforms and techniques and applications involving integrating single-cell sequencing and optical measurements at high throughput. The microfluidics tools that will be discussed in this Dissertation aim to be a platform for performing single-cell multi-parameter and multi-omics techniques that will help further our understanding of cellular identity and how genotype informs phenotype at the single-cell level.

Table of Contents

List of Figures

List of Tables

Chapter 1 : Introduction

1.1 The single cell as the fundamental unit of biology

All of biology is comprised of cells. While the simplest organisms comprise a single cell, multicellular organisms organize cells into organs and organ systems. Despite containing the same genetic information, cells specialized for motor movement have vastly different function than those involved with digestion. Specialization in cellular function is one of enabling factors for the evolution from unicellular organisms to much larger and complex multicellular organisms. Because not all organisms are alike, a key question of biology is understanding the driving factors behind this organism to organism heterogeneity, especially within members of the same species.

The first observed form of heterogeneity was phenotypic heterogeneity, which refers to the slight differences between individuals of the same species. These differences may appear outwardly obvious, such as differences in size, or reflected in habits, such as preferred sleeping hours. Pioneering geneticists were the first to understand that this phenotypic heterogeneity has its roots within the genetic code. At the single cell level, a similar question can be asked: are differences in cellular function attributed to differences in how cells interpret their DNA? Sequencing-based assays developed over the past two decades have helped researchers elucidate this question. The resolution of these sequencing-based assays has improved to the point where it is now possible to sequence genetic material deriving from a single cell. However, one general limitation of these assays is the ability to connect sequencing-based measurements to other measurements of cellular phenotype. To build a more robust understanding of cellular

heterogeneity, we need a new series of tools to further explore how genotype and phenotype are connected at the single cell level.

1.2 Probing cellular heterogeneity through the lens of a microscope

The term 'cell' was first coined by Robert Hooke in the 17th century, when he remarked how the walls separating different cells in a slice of cork looked like cellula, the chambers in which monks resided (1). Key in enabling Hooke to make this observation was the microscope, which was also invented in the 17th century. Another scientist working during this period, Antonie van Leeuwenheok, used the microscope to observe microbes, describing different species of protists and bacteria in his seminal article 'Letter on the protozoa" (2). He and other scientists noted subtle differences between cells of the same species, thus making the microscope the first tool that scientists used to directly study cellular heterogeneity.

The microscope remains today a key tool in investigating cellular heterogeneity. Today's microscopes can perform imaging with sufficient resolution to resolve sub-cellular structures and in 3-D. Fluorescence labeling of specific targets within cells, whether they be surface epitopes, intracellular proteins or nucleic acids is also a key aspect of modern microscopy. Our understanding of cellular heterogeneity has been shaped by what we can visualize with microscopes: features such as cell size and shape, the expression of specific surface receptors, and the types of cells which neighbor a given cell. Super resolution microscopy techniques developed in the past two decades now enable the direct visualization of the statics and dynamics of single proteins, transcripts, and other molecules within cells (3). Our understanding of cellular heterogeneity began with the microscope; advances in imaging capabilities will ensure that the microscope will remain a key fixture in single-cell biology in the future.

1.3 Probing cellular heterogeneity with sequencing

While microscopes have largely been used to study phenotypic aspects of cells, the development of DNA sequencing enabled scientists to investigate the genetic basis of cellular heterogeneity. DNA is the origin of all the sub-cellular structures which come together to form a complete cell; this statement is understood as the Central Dogma, describing information flow in a cell starting from genomic DNA, from which select portions are transcribed into RNA and then translated into protein (4). The first applications of DNA sequencing were towards sequencing of complete genomes, starting from viruses and prokaryotic organisms, which have much smaller genome sizes than eukaryotes. The complete sequencing of the human genome in 2001 (5) culminated a decades' worth of effort. Since then, the development of next-generation sequencing (NGS) has increased our sequencing output by multiple orders of magnitude. At the same time, the sequencing costs per million bases has decreased thanks to NGS to the point where they are accessible to most research groups.

NGS has made feasible sequencing-based assays to probing entire -omics layers within cells. These approaches are unbiased and label-free, taking advantage of the increased sequencing capacity possible by NGS to capture entire classes of nucleic acids within cells. One of methods develop as a result was mRNA-sequencing (RNA-seq), which enabled researchers to take a snapshot of the entire transcriptome of cells. In RNA-seq, a poly-T capture probe can capture all mRNA within a cell due to the monolithic structure of eukaryotic mRNA which terminates in a 3' polyadenine tail (6). Following mRNA capture, reverse transcription can be used to convert the RNA into DNA, which can be processed and sequenced on a NGS machine. The input material to a RNA-seq study gradually has been scaled down to the point where single cells can be studied (7). Single-cell RNA-seq (scRNA-seq) has been an important technique in

advancing our notion of cellular heterogeneity (8). scRNA-seq demonstrated variations in cell-to-cell patterns of gene expression that would have otherwise been hidden by a bulk sequencing approach (9). The classical view of cells being grouped into broad classifications has evolved to become one where each cell within an organism has a unique identity. This identity is dynamic and can be shaped by such factors as spatial location, neighboring cells, and fluctuations in cellular state.

The throughput of scRNA-seq platforms has increased dramatically in the past decade, with it now being possible to sequence over 10,000 cells per experiment (10). Still, the transcriptome is just one -omics layer within the cell. Parallel to the development of RNA-seq were the development of assays studying -omics layers within cells. There now exist a host of methods sequencing the epigenetic landscape within cells, looking at aspects such as DNA associated with specific nuclear proteins (11), chromatin accessibility (12), and chromatin modifications (13). The development of this plethora of single-cell sequencing tools has contributed to a refined understanding of cellular heterogeneity that moves beyond what is visible under a microscope.

1.4 Droplet microfluidics facilitates high throughput single-cell sequencing

A critical step in any single-cell sequencing workflow is the isolation of individual cells. The first demonstrations of scRNA-seq involved isolating individual cells into centrifuge tubes or wells of a 96 well plate (7,14). While commercially available instruments like flow cytometers can automate the sorting of single cells into well plates, this approach lacks scalability. To analyze thousands of cells, many individual well plates would be required, leading to inefficiencies in reagent cost and experimental time. A large contributing factor this inefficiency is a length scale mismatch between cells and microwell plates: the size of a single cell is on the

order of ~10 um, and each well of a 96-well plate is ~9000 um in diameter. Ideally, single cells should be isolated into reaction vessels which more closely match their size.

Droplet microfluidics is a sub-field of microfluidics which concerns the manipulation of emulsions of sub-nanoliter droplets. Cells, solid polymeric particles, and reagents can all be packaged within these droplets to perform various bioassays. As these sub-nanoliter droplets are between 10-100 um in diameter, they make an effective reaction vessel for performing high throughput single-cell assays, as their size is on the same order of magnitude as that of a cell. One of the common methods for generating droplet emulsions involves merging two flowing streams of aqueous and oil solutions at a flow focusing junction (**Figure 1-1**) (15). At the junction, the competition of the interfacial tension of the aqueous phase and the shear stress due to laminar flow mediate the regular pinching of aqueous droplets in a process known as the dripping mode of droplet generation (16,17). The capillary number (Ca), defined as the ratio of the product of the velocity and viscosity of a fluid to its surface tension, is the main nondimensional constant mediating the dripping regime (17). While other modes of droplet generation exist at low and high Ca (16), the dripping regime at moderate Ca is the most common.

Droplet-based assays are high throughput and scalable. Each droplet can be treated as its own reaction vessel as it is isolated from other droplets in an emulsion due by the presence of a stabilizing surfactant layer. Ultrahigh throughput assays utilizing droplet emulsions have been

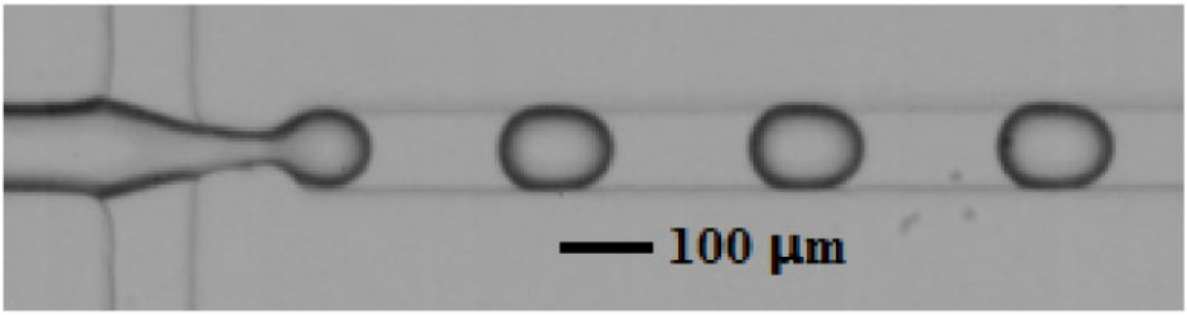

Figure 1-1. Droplet generation at a flow-focusing junction.
An aqueous stream meets two orthogonal oil streams. Adapted from (15).

applied to digital PCR-based detection (18), large-scale enzymatic function screens (19), and drug screens (20). Droplets can be used to efficiently isolate single cells. The encapsulation rate of cells and other solid particles in droplets follows Poisson statistics (21). To achieve accurate single-cell isolation, the input cell suspension is diluted such that upwards of 95% of generated droplets will not contain any cell. As it is facile to generate emulsions consisting millions of droplets, many thousands of cells can also be readily isolated within droplets. Droplet microfluidics has been applied broadly within the high-throughput single-cell sequencing space, in applications such as targeted genomic loci (22), mRNA (23), and epigenome sequencing (24).

A major challenge of adapting any bioassay to droplets is that unlike with wells in a well-plate, droplets cannot be directly manipulated. While microfluidic devices have been engineered to divide, combine, and selectively enrich droplet emulsions (25,26,27), they add additional complexity to droplet-based workflows. To minimize the number of assay steps performed within droplets, in several droplet-based single-cell sequencing workflows, following isolation and lysis of single cells within droplets, all of the genomic material in each droplet is labeled

with a drop-specific oligonucleotide sequence (23,28). This process, known as droplet barcoding, enables pooling of all droplets during downstream reaction steps and subsequent sequencing library preparation in bulk solution. Data originating from individual droplets can be isolated following sequencing by bioinformatic grouping by the drop-specific barcode. In droplet-based scRNA-seq workflows, a common strategy is to use barcoded poly-T capture beads (23,29); once a cell and bead are co-encapsulated within a drop, the cell is lysed and all of its mRNA is barcoded. Droplet barcoding is key to enabling the development of single-cell sequencing experiments with throughput in the tens of thousands of cells per experiment.

1.5 Towards single-cell muti-omics

A cell's phenotype is derived from the various -omics layers which describe the interpretation of genomic DNA into functional macromolecules. While large-scale studies of transcriptomics at the single-cell level have been achieved up to the whole-organism level (30), to obtain a richer understanding of how each cell within an organism differs from the others, we need to be able to integrate information from other -omics layers. Progress has been made towards single-cell multi-omics, with methods now able to simultaneously probe DNA and RNA (31,32) but these methods largely lack the scalability demonstrated by droplet-based single-cell sequencing platforms due to the increased complexity of these workflows. Furthermore, there are a limited number of techniques available for directly correlating an image-based measurement of a cell's phenotype with a sequencing-based measurement (33,34). In these techniques there is also a tradeoff between the throughput and information output. To be able to further understand how a phenotype and genotype are connected at the single-cell level, we need a new suite of tools to enable both high-throughput and comprehensive measurement of single cells.

In this dissertation, I summarize progress I have made towards developing a new suite of high-throughput tools linking single-cell sequencing and imaging. In Chapter 2, I describe a new droplet microfluidic device developed in the Abate and Gartner labs. In Chapter 3, I apply the device presented in Chapter 2 to perform high-throughput fluorescence and RNA-seq measurements of single cells. In Chapter 4, I describe a novel strategy for making droplet emulsions for applications such as single-cell sequencing with droplet barcoding. Lastly, I close in Chapter 5 with detailing the steps I have made towards linking microscopy and RNA-seq measurements at single-cell scale.

1.6 References

1. Ellinger I, Ellinger A. Smallest unit of life: Cell biology. In Ellinger I, Ellinger A. Comparative Medicine: Anatomy and Physiology.: Springer-Verlag Wien; 2014. p. 19-33.

2. Lane N. The unseen world: reflections on Leeuwenhoek (1677) 'Concerning little animals'. Philosophical Transactions of the Royal Society B: Biological Sciences. 2015 Apr 19; 370(1666): 20140344.

3. Leung BO, Chou KC. Review of super-resolution fluorescence microscopy for biology. Applied spectroscopy. 2011 Sep 1; 65(9): 967-80.

4. Crick F. Central dogma of molecular biology. Nature. 1970; 227(5258): 561-563.

5. Craig Venter J, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The sequence of the human genome. Science. 2001 Feb 16; 291(5507): 1304-1351.

6. Wang Z, Gerstein M, Snyder M. RNA-Seq: A revolutionary tool for transcriptomics. 2009 Jan..

7. Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, et al. mRNA-Seq whole-transcriptome analysis of a single cell. Nature Methods. 2009 Apr 6; 6(5): 377-382.

8. Wagner A, Regev A, Yosef N. Revealing the vectors of cellular identity with single-cell genomics. 2016 Nov 8..

9. Shalek AK, Satija R, Adiconis X, Gertner RS, Gaublomme JT, Raychowdhury R, et al. Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. Nature. 2013 May 19; 498(7453): 236-240.

10. Svensson V, Vento-Tormo R, Teichmann SA. Exponential scaling of single-cell RNA-seq in the past decade. 2018 Apr 1..

11. Rotem A, Ram O, Shoresh N, Sperling RA, Goren A, Weitz DA, et al. Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. Nature Biotechnology. 2015 Nov 1; 33(11): 1165-1172.

12. Buenrostro JD, Wu B, Litzenburger UM, Ruff D, Gonzales ML, Snyder MP, et al. Single-cell chromatin accessibility reveals principles of regulatory variation. Nature. 2015 Jul 23; 523(7561): 486-490.

13. Luo C, Rivkin A, Zhou J, Sandoval JP, Kurihara L, Lucero J, et al. Robust single-cell DNA methylome profiling with snmC-seq2. Nature Communications. 2018 Dec 1; 9(1): 1-6.

14. Picelli S, Faridani OR, Björklund ÅK, Winberg G, Sagasser S, Sandberg R. Full-length RNA-seq from single cells using Smart-seq2. Nature Protocols. 2014 Jan 2; 9(1): 171-181.

15. Zeng W, Xiang D, Fu H. Prediction of Droplet Production Speed by Measuring the Droplet Spacing Fluctuations in a Flow-Focusing Microdroplet Generator. Micromachines. 2019 Nov 25; 10(12): 812.

16. Romero PA, Abate AR. Flow focusing geometry generates droplets through a plug and squeeze mechanism. Lab on a Chip. 2012 Dec 21; 12(24): 5130-5132.

17. Xu JH, Li SW, Tan J, Luo GS. Correlations of droplet formation in T-junction microfluidic devices: From squeezing to dripping. Microfluidics and Nanofluidics. 2008 May 30; 5(6): 711-717.

18. Hindson BJ, Ness KD, Masquelier DA, Belgrader P, Heredia NJ, Makarewicz AJ, et al. High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. Analytical Chemistry. 2011 Nov 15; 83(22): 8604-8610.

19. Romero PA, Tran TM, Abate AR. Dissecting enzyme function with microfluidic-based deep mutational scanning. Proceedings of the National Academy of Sciences of the United States of America. 2015 Jun 9; 112(23): 7159-7164.

20. Brouzes E, Medkova M, Savenelli N, Marran D, Twardowski M, Hutchison JB, et al. Droplet microfluidic technology for single-cell high-throughput screening. Proceedings of the National Academy of Sciences of the United States of America. 2009 Aug 25; 106(34): 14195-14200.

21. Collins DJ, Neild A, deMello A, Liu AQ, Ai Y. The Poisson distribution and beyond: Methods for microfluidic droplet production and single cell encapsulation. 2015 Jul 21..

22. Pellegrino M, Sciambi A, Treusch S, Durruthy-Durruthy R, Gokhale K, Jacob J, et al. High-throughput single-cell DNA sequencing of acute myeloid leukemia tumors with droplet microfluidics. Genome Research. 2018 Sep 1; 28(9): 1345-1352.

23. Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. Cell. 2015 May 30; 161(5): 1202-1214.

24. Grosselin K, Durand A, Marsolier J, Poitou A, Marangoni E, Nemati F, et al. High-throughput single-cell ChIP-seq identifies heterogeneity of chromatin states in breast cancer. Nature Genetics. 2019 Jun 1; 51(6): 1060-1066.

25. Abate AR, Weitz DA. Faster multiple emulsification with drop splitting. Lab on a Chip. 2011 Jun 7; 11(11): 1911-1915.

26. Niu X, Gulati S, Edel JB, Demello AJ. Pillar-induced droplet merging in microfluidic circuits. Lab on a Chip. 2008 Nov 1; 8(11): 1837-1841.

27. Sciambi A, Abate AR. Accurate microfluidic sorting of droplets at 30 kHz. Lab on a Chip. 2015 Jan 7; 15(1): 47-51.

28. Lan F, Demaree B, Ahmed N, Abate AR. Single-cell genome sequencing at ultra-high-throughput with microfluidic droplet barcoding. Nature Biotechnology. 2017 Jul 1; 35(7): 640-646.

29. Klein AM, Mazutis L, Akartuna I, Tallapragada N, Veres A, Li V, et al. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. Cell. 2015 May 30; 161(5): 1187-1201.

30. Han X, Wang R, Zhou Y, Fei L, Sun H, Lai S, et al. Mapping the Mouse Cell Atlas by Microwell-Seq. Cell. 2018 Feb 22; 172(5): 1091-1107.e17.

31. Macaulay IC, Haerty W, Kumar P, Li YI, Hu TX, Teng MJ, et al. G&T-seq: Parallel sequencing of single-cell genomes and transcriptomes. Nature Methods. 2015 May 28; 12(6): 519-522.

32. Hou Y, Guo H, Cao C, Li X, Hu B, Zhu P, et al. Single-cell triple omics sequencing reveals genetic, epigenetic, and transcriptomic heterogeneity in hepatocellular carcinomas. Cell Research. 2016 Mar 1; 26(3): 304-319.

33. Lane K, Van Valen D, DeFelice MM, Macklin DN, Kudo T, Jaimovich A, et al. Measuring Signaling and RNA-Seq in the Same Cell Links Gene Expression to Dynamic Patterns of NF-κB Activation. Cell Systems. 2017 Apr 26; 4(4): 458-469.e5.

34. Yuan J, Sheng J, Sims PA. SCOPE-Seq: A scalable technology for linking live cell imaging and single-cell RNA sequencing. Genome Biology. 2018 Dec 24; 19(1): 1-5.

Chapter 2 : Printed Droplet Microfluidics for on demand printing of picoliter droplets and cells

Elements of the following chapter are reprinted from the manuscript "Printed droplet microfluidics for on demand dispensing of picoliter droplets and cells," by Russell H. Cole, Shi-Yang Tang, Christian A. Siltanen, Payam Shahi, Jesse Q. Zhang, Sean Poust, Zev J. Gartner, and Adam R. Abate, published in *Proceedings of the National Academy of Sciences of the United States of America* on August 15, 2017, Vol. 114, Issue 33, pages 8728-8733.

2.1 Abstract

Although the elementary unit of biology is the cell, high-throughput methods for the microscale manipulation of cells and reagents are limited. The existing options either are slow, lack single-cell specificity, or use fluid volumes out of scale with those of cells. Here we present printed droplet microfluidics, a technology to dispense picoliter droplets and cells with deterministic control. The core technology is a fluorescence-activated droplet sorter coupled to a specialized substrate that together act as a picoliter droplet and single-cell printer, enabling high-throughput generation of intricate arrays of droplets, cells, and microparticles. Printed droplet microfluidics provides a programmable and robust technology to construct arrays of defined cell and reagent combinations and to integrate multiple measurement modalities together in a single assay.

2.2 Introduction

While the basic building block of life is the cell, humans comprise trillions of molecularly distinct cells whose interactions give rise to the functional properties of our bodies. A more detailed understanding of this network of interacting cells is required to reveal the

etiology of human disease and to enable the regenerative medicines of the future, yet its complexity defies detailed study *in situ*. However, current techniques for cell manipulation, isolation, and combination are rudimentary and of limited value for high throughput studies, particularly when multiple cells of different type are required for the phenomena of interest.

Cell suspensions can be loaded into microfluidic droplets or wells to isolate large numbers of cells, yet the uncontrolled and statistical nature of cell loading leads to many volumes that are unusable because they are empty or contain undesired cellular contents. Deterministic dispensing to well plates with fluorescence-activated cell sorting is an efficient single cell handling technique, though depositing cells into microliter volumes, precludes many advantages of assays in small volumes, complicating the formation of direct cellular contacts, and diluting secreted molecules to non-physiological concentrations. Tackling problems with increased cellular complexity requires that biologically-meaningful experiments be carefully constructed and performed with massive reductions in reactor volume and significant increases in throughput.

In this paper, we address this fundamental limitation of current technologies using Printed Droplet Microfluidics (PDM), a new kind of microfluidics that prints droplets containing reagents and cells to defined arrays, where they can be deterministically combined and manipulated. The core of the approach is a fluorescence-activated droplet sorter (1) acting as an intelligent cell and droplet dispenser coupled with a motorized stage. Candidate droplets are cycled through the instrument and scanned to determine their contents, then printed to a substrate if a droplet contains the reagent or cell required for a given position in the array. In this way, droplet printing is made deterministic by intelligent sorting, allowing the construction of cell and reagent arrays of unprecedented intricacy. PDM combines the control and generality of

microliter pipetting in well plates, the precise single cell dispensing of FACS, and the scalability and single cell sensitivity of droplet microfluidics.

2.3 Results

2.3.1 A microfluidic robot for droplets and single cell dispensing.

Microliter pipetting has become a universal tool in biology due to its ability to deliver precise volumes of multiple reagents into well plate reactors for subsequent assays. These tools are easily reconfigurable using software, allowing the same instrument to be used in applications as diverse as studying cancer cells to engineering bacteria with unnatural properties (2, 3), and are now broadly used across the biomedical sciences.

Microliter fluid handlers, however, retain two major functional limitations: the microliter volumes required limit the number of reactions that can be performed; and the methods for dispensing liquids are not optimized to deliver precise numbers of individual particles or cells. To increase throughput, volume must scale down, which is the core concept of droplet microfluidics: a field that studies reactions in picoliter droplets six orders of magnitude smaller than fluid volumes typical in microtiter plates. By operating at the scale of individual cells, droplet microfluidics affords massive increases in speed and reaction number compared to microliter pipetting. Additionally, by combining droplet microfluidics with an in line fluorescence-activated droplet sorter, it becomes possible to insure that every droplet contains one and only one cell. However, droplet microfluidics remains an incomplete solution for picoliter fluid handling, even when combined with in line droplet sorters. Specifically, the process of forming, combining, sorting, and analyzing picoliter droplets requires custom microfluidic devices unique to each task. Therefore, each new task requires a new device and an expert to design, build, and operate it. To impact biology with better fluid handling, a new

approach is required that achieves the throughput and scalability of droplet microfluidics, while retaining the flexibility and reconfigurability of microliter fluid handlers.

Several technologies have been developed for the controlled and flexible manipulation of microfluidic droplets. Digital microfluidics utilizes electrowetting to move, combine, and split nanoliter-scale droplets on a specialized substrates, and have been used to prepare samples for sequencing (4) and to probe inter-cellular signaling on the single cell level (5). A number of approaches have used open-ended microfluidic devices to dispense nL volumes to substrates(6, 7). Though these technologies enable droplet microfluidics in arrayed formats, the ability to manipulate single cells and rapidly switch reagents is lacking.

We have therefore developed a new microfluidics technique that further breaks through the constraints of conventional droplet microfluidics. Our method, PDM, miniaturizes array pipetting to the picoliter scale. The core of the technology is a microfluidic sorting-on-demand device that acts as a deterministic picoliter droplet dispenser on a motorized substrate (Fig. 2-1). The key attribute of this device is that it can select a specific droplet from a set of candidates and dispense it to a substrate (Fig. 2-2a, device schematic, Fig. 2-2b). It does this in a completely deterministic and programmable fashion, allowing the printing of intricately defined arrays of

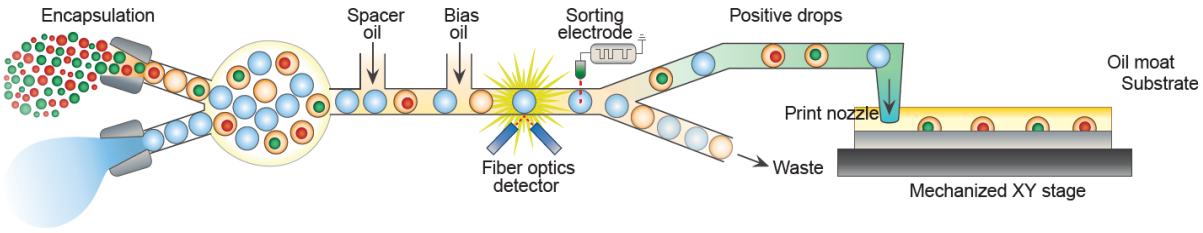

Figure 2-1. Printed Droplet Microfluidics workflow.
The printer consists of a sorter-based print head that selectively dispenses microfluidic droplets to a substrate under a cover of oil. The printing process is automated by coupling the print head to a motorized stage. Deterministic combinations of cells and reagents can be formed at each location through additive droplet printing.

picoliter to nanoliter volumes based on a software encoded print file. Moreover, it is fast, selecting from a set of candidate droplets at up to kilohertz frequencies (8).

To implement a microfluidic sorter as a droplet printer, we overcame several technical challenges. Foremost, droplet sorting requires oil as a carrier phase that must be displaced for the droplets to reach the substrate. To overcome the viscous and buoyancy forces that would slow or prevent deposition of droplets, we require a counter force to pull the droplets to the print locations. A simple and general technique for this is dielectrophoresis, commonly used to pull conductive droplets in oil for high speed sorting (1). To implement dielectrophoretic trapping, we construct a substrate consisting of an array of bipolar electrodes under a thin sheet of dielectric (Fig. 2-2c). When energized, the electrodes emit field extending above the substrate, causing droplets to settle into nanowells patterned on top of the dielectrophoretic traps where they are immobilized (Fig. 2-2d).

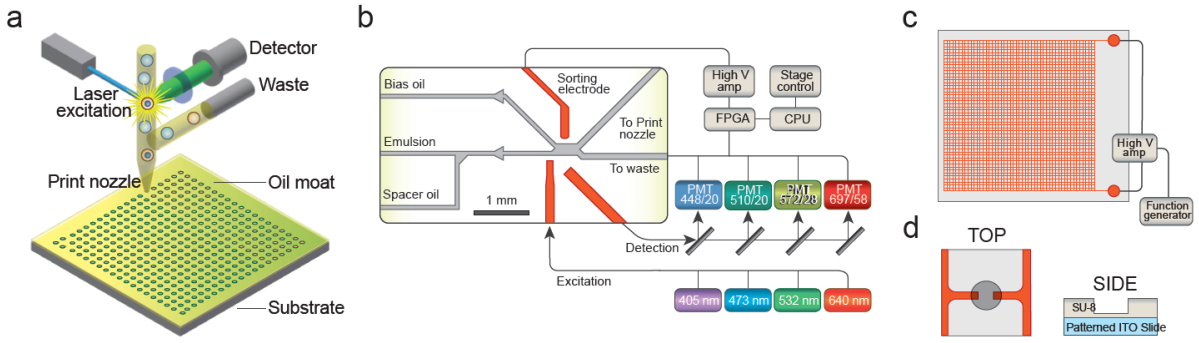

Figure 2-2 Components of the PDM system.
a) Droplets flow out of the nozzle in carrier oil to nanowell features on the substrate, where they are immobilized. b) The print head consists of a dielectrophoretic droplet sorter fitted with fiber optics. Four color excitation and detection sort droplets of interest to the print nozzle while all other droplets flow into the waste stream. Droplet printing can utilize on-chip drop making or the reinjection of pre-formed emulsions. c). The printing substrate consists of a nanowell array fabricated on top of a grid of dielectrophoretic traps. Typical nanowells in this work have a diameter of 150 µm and a depth of 75 µm, and are fabricated on a 400 µm pitch. d) Wells are photo-patterned from SU8 on top of a laser etched ITO-coated glass slide.

2.3.2 Fundamental fluidic operations

Single droplet delivery is accomplished by positioning the outlet nozzle of the sorter above a dielectrophoretic target, dispensing a single droplet by activating a dielectrophoretic field in the print head to divert a droplet to the exit channel. The droplet is pulled into the trap minimum, displacing the oil as it reaches, and touches, the substrate (Fig. 2-3a). The substrate electrodes are covered by an insulating film to prevent current flow within an immobilized droplet.

A persistent challenge in droplet microfluidics is that it is difficult to add many reagents to pre-existing droplets. Current methods, such as coalescence (9) and picoinjection (10) are effective for a limited number of additions, and not well suited for the multitude of workflows that require reagent additions to occur at different times. In PDM addition can be accomplished by hovering over a printing position and dispensing additional droplets (Fig. 2-3b). Each dispensed droplet is pulled to the trap minimum, where it immediately

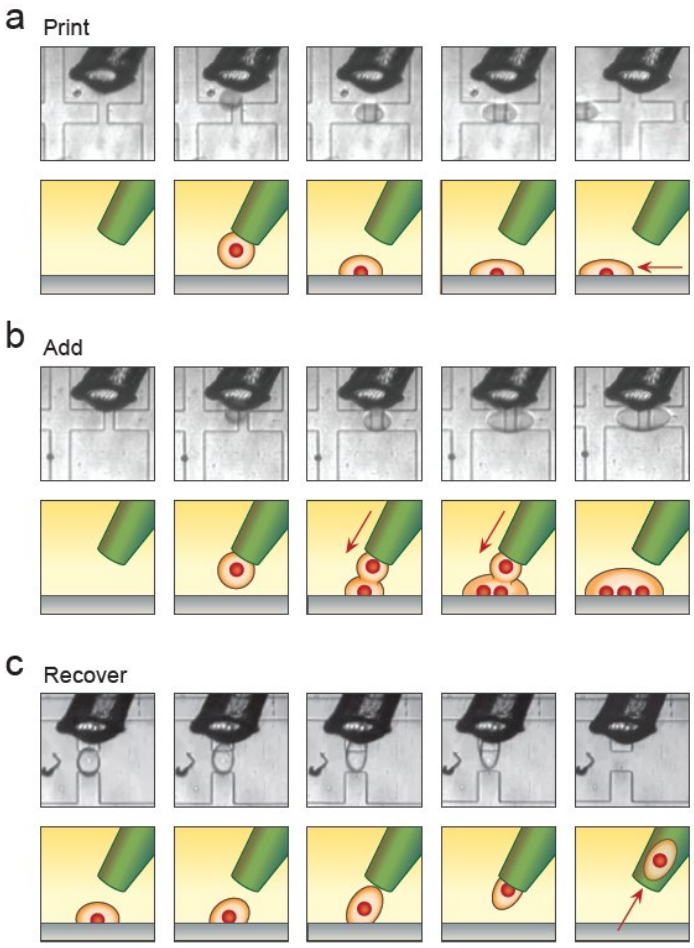

Figure 2-3. Fundamental droplet printer operations performed on a flat substrate.
a) Print. Deposition of a single droplet at an array position, after which the printer rasters to the next position. b) Add. Successive printing and merging of three consecutive droplets to the same position. c) Recover. Removal of a single droplet from the substrate via suction.

electrocoalesces with the droplet already positioned there. The ease with which PDM can perform deterministic and programmable addition of reagents to preexisting droplets makes it possible to perform a suite of manipulations that are too impractical to reliably perform with flowing droplet microfluidics.

Droplet dispensing and addition are the basic operations for generating arrays with defined composition. Often, the next step in an experiment is to image the array and, in some cases, recover select droplets for further study, such as additional rounds of culture or sequencing. Recovery can be accomplished by simply vacuuming the droplet from the substrate, even with the electric field on (Fig. 2-3c). Recovery of mass numbers of droplets from a printed array via suction requires that a suction nozzle raster over the entire print substrate. With the current hardware, indiscriminate recovery of droplets from the array can be performed at ~4 Hz, and is essentially limited by the rate of stage movement. Recovery of contents from individual positions and placement of these fluid volumes into specific wells in a well plate is slower to allow for the transit of individual droplets to a unique external location and can be performed at ~0.1 Hz.

2.3.3 Programmed printing of intricate droplet arrays.

Microliter pipetting is used universally in the biomedical sciences due to its flexibility and the control with which it can combine multiple reagents. Picoliter dispensing technologies such as PDM must thus be able to create arrays of comparable definition and complexity to ensure similarly broad utility. Competing technologies for scaling-down reagent transfers rely on inkjet or acoustic dispensing (11), yet these technologies are used to construct experiments at microliter scales to prevent rapid droplet dehydration. Furthermore, constructing combinatorial experiments requires either additional printing channels, or a time-consuming reagent switchover

step, a proposition that becomes onerous for highly combinatorial screens. By contrast, PDM switches elegantly between reagents contained in a mixed emulsion "ink" by dielectrophoretic droplet sorting, possible at kHz rates (8).

The current system print rate is related to three relevant delay times: the wait time for a desired droplet to enter the detection region of the print head, the droplet travel time from the print head sorting junction to the printing substrate, and the time for the stage to move from one nanowell array position to the next. For the system described here, droplets are typically generated or reinjected at 50-100 Hz, leading to wait times of 10-20 mS between uniform dye droplets and 0.2-0.4 S between droplets encapsulating single cells (at 1:20 limiting dilution). The travel time for a droplet to reach the substrate is ~0.1 S, and the mechanical stage can move between positions 400 μm apart in a 0.25 S. Our control algorithm enables droplets to be sorted with a fixed time delay, so that the 0.1 S of droplet travel occurs during stage movement, and does not limit the maximum 4 Hz print rate of the system when droplets delay times are short. Because multiple droplets can be sorted and en route to the substrate at a given time, printing multiple dye droplets to a position requires that the print nozzle hover over the print substrate for the additional sorting time associated with these droplets. For cell experiments in this work, we chose extremely selective gates to eliminate the possibility of sorting cell doublets, which in turn limited the availability of desired cells to ~1 Hz. Although beyond the scope of the proof of concept experiments demonstrated in this work, increasing printing throughput requires several straightforward system modifications. Our lab has demonstrated dielectrophoretic droplet sorting at rates of up to 30 kHz (8), so detecting and sorting droplets at 10-100X greater frequencies is readily achievable by increasing oil and reagent flows. Similarly, the time scale for droplet travel to the substrate is reducible by 10-100X through increased flow and decreased channel

dimensions. The use of a faster piezo stage will enable the print head to move between 400 µm array positions at ~ 20 Hz, and address 10,000 position arrays in ~ 10 minutes.

Sorting-on-demand requires that the contents of each candidate droplet be identifiable via in-flow fluorescence in the microfluidic sorter. Since most reagents of interest are not fluorescent, we therefore include "fluorescence barcodes" in the droplets. For example, 100 different reagent solutions can be labeled by mixing two dyes at ten concentrations. Such combinatorial labeling has been demonstrated using lanthanide nanophosphors for over 1000 unique spectral barcodes, indicating that reagent labeling is not a meaningful limitation (12).

To demonstrate PDM's ability to construct intricate arrays, we generate arrays of defined composition from two solutions of fluorescent dye, green (FITC) and red (TRITC) (Fig. 2-4). The dye solutions are injected via syringe pumps and encapsulated into droplets in a double T-junction upstream of the sorter (Fig. 2-4a, *inset*). The device forms droplets of both solutions in an alternating sequence, which is used to print the desired combination via sorting-on-demand. The droplets are scanned for fluorescence using integrated fiber optics, appearing as two tight clusters on a red versus green scatter plot (Fig. 2-4a). These clusters define "red" or "green" population gates, allowing the instrument to print them in a deterministic fashion following the print file.

The simplest printing pattern is an array of identical droplets with identical contents. To print this array, we program 4 green droplets to be printed into every array position in the print file, and print the file. The instrument follows the program, sorting 4 green droplets into every spot on the array, yielding a uniform population of 100 green droplets, each composed of 4 combined droplets, and each immobilized in a nanowell structure located above a dielectrophoretic trap (Fig. 2-4b). Uniform sets of microscale reactors like this, however, can be

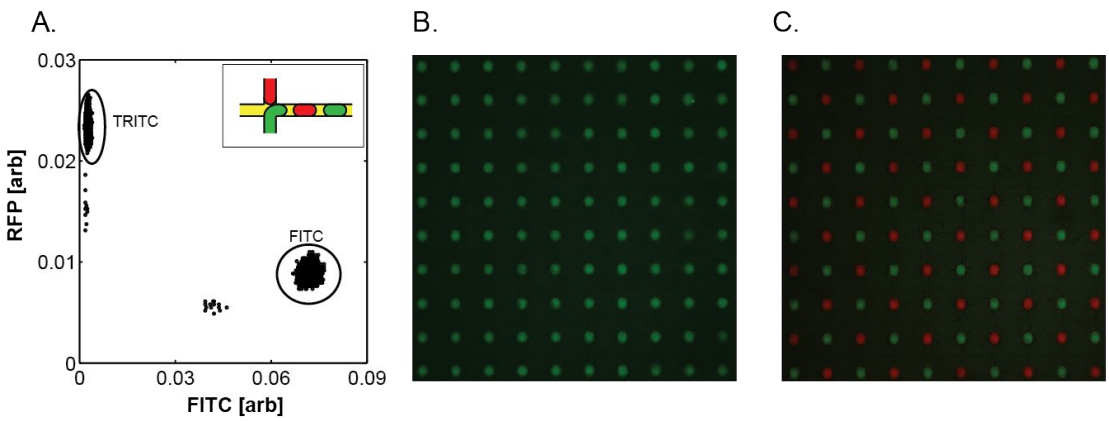

Figure 2-4. Fluorescent dye printing.
a) Flow cytometry data for print head utilizing a dual drop maker to produce emulsions on chip. b) 100 position array printed with 4 droplets of FITC at each position from the emulsion shown in A. c) 100 position array with alternating positions printed with 4 droplets of either FITC or TRITC.

created with many methods, including common microfluidic droplet generation, passive filling of nanoliter wells, or slip-chip systems (13–15). A more complex pattern that is impossible to build with these methods is alternating red and green droplets. To construct this array, we program the print file appropriately, and print it. The instrument sorts the droplets on demand in accordance with the file, generating the desired alternating array (Fig. 2-4c). In this way, changing the experiment performed by the instrument requires only re-writing the print file, rather than changing hardware – a departure from conventional flowing droplet microfluidics.

Encapsulating reagents on the device using a double T-junction is simple and allows printing of two solutions. However, this becomes impractical when many reagents must be printed, because this would require many feed lines connected to the tiny ~1 cm^3 sorting device. A better approach is to emulsify all solutions, combine them in a mixed emulsion, and reinject this mixed emulsion "ink" through a single line. To illustrate this, we generate a mixed emulsion of four droplet types, labeling them with unique combinations of FITC, TRITC, Cy5, and Cascade Blue (CB) dyes. The detection scatter plot has four clusters, one for each droplet type

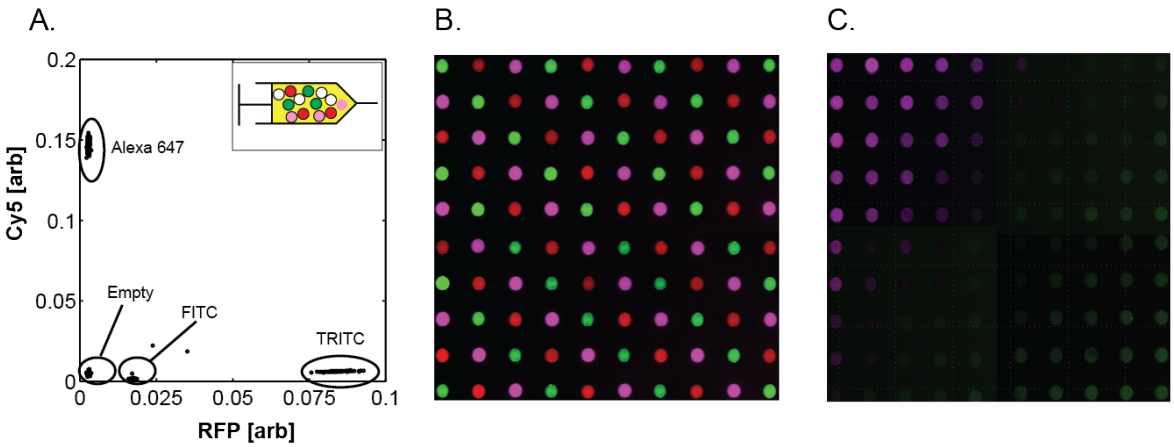

Figure 2-5. Printing of three fluorescent dyes.
a) Flow cytometry data from a reinjected emulsion containing 4 dyed droplet types. (empty, FITC, TRITC, and Cy5 (Alexa 647)) b) 100 position array printed with alternating positions containing 4 droplets of FITC, TRITC, and Cy5. c) position well gradient array printed with 8 droplets containing combinations of FITC, Cy5, and empty droplets.

(Fig. 2-5a). We use these to define color gates and load a print file with a pattern that prints 4 droplets of 3 different types to each position in a organized mixed color array (Fig. 2-5b).

Arrays like these illustrate the potential for using PDM to design and execute totally independent reactions at each spot on an array. Often, however, the objective will be to scan a parameter space that depends on the relative concentrations of multiple reagents. This can be accomplished by systematically varying combinations across the array (Fig. 2-5c). A major advantage of PDM over flowing droplets is that grid locations relate the composition of every droplet in the printed array; the array is generated in a deterministic fashion using a print file that contains all information about what was combined at each spot. While in flowing droplet microfluidics it is possible to generate varying concentrations of reagents using oscillating flows or random droplet merger (16, 17), it is far more difficult to determine what was combined in each droplet afterwards, since they are not neatly stored on an array.

Another key advantage of PDM is the ability to perform experiments at high throughput. To demonstrate this, we use green droplets to print an image file to a 10,000-position array (Fig.

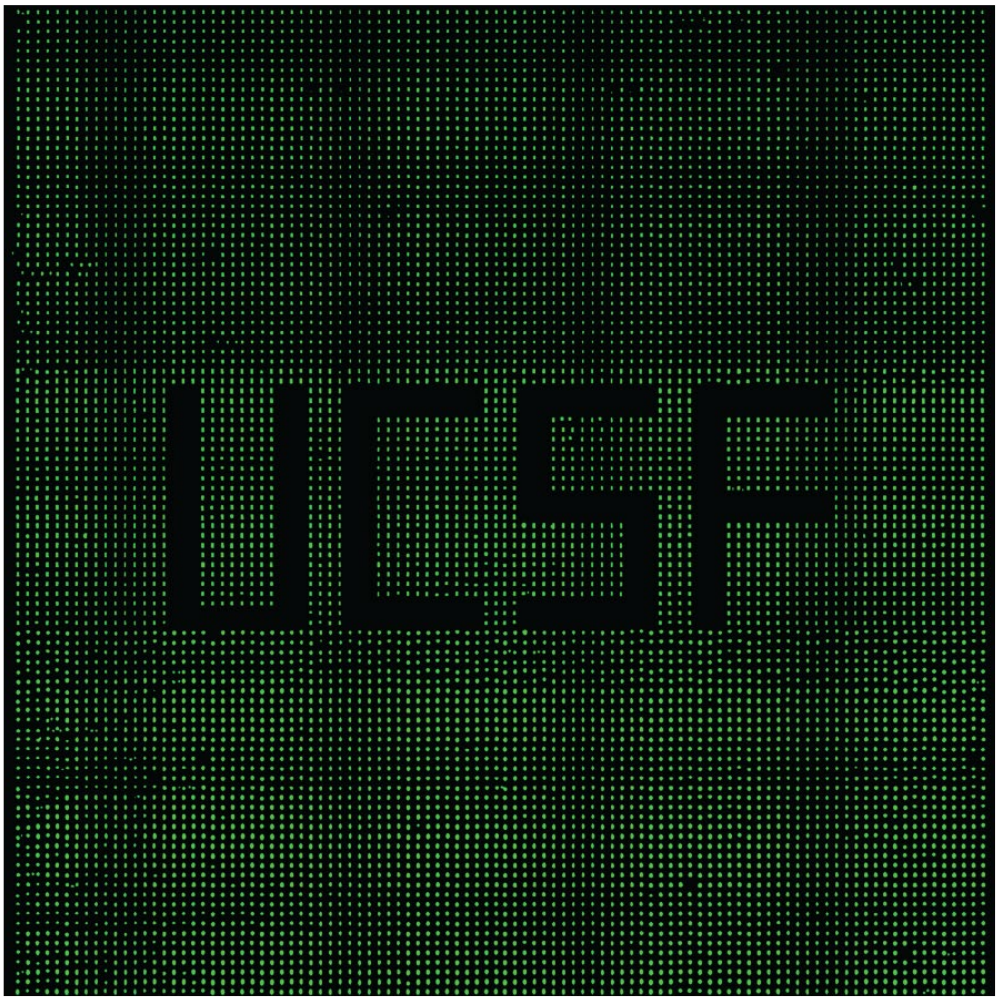

Figure 2-6. Construction of a binary image by printing FITC droplets to a 10,000 position array.

2-6). The image is stitched from 3 separate printing runs that each populate a subregion of the total array. The print rate for the array is 1.5 Hz, leading to an elapsed printing time of ~2 hours for the entire array. This demonstrates that the number of experiments that can be constructed on a single substrate using PDM (10^4) is two orders of magnitude larger than the typical scales used in well plate fluid handling (10^2).

2.3.4 On-demand printing of single cells and complex cell combinations.

Biological systems comprise populations of heterogeneous cells that are important to the form and function of the system as a whole. While there are now good methods for high throughput analysis of single cells, fewer methods exist for constructing and studying cell combinations. This would be important, for example, for reconstituting cellular interactions *in vitro* to study the impact of cellular heterogeneity on tissue function. Additionally, combinations of cells could be used for studying tissue self-organization, immune-tumor interactions, or the effects of tumor heterogeneity on drug response. Methods based on random encapsulation in droplets or chambers do not allow generation of defined cell combinations (18).

PDM provides an elegant solution to generating any desired combination of cells, by deterministically and repeatedly printing single cells of different type to droplet arrays. To illustrate this, we use the instrument to generate defined combinations of green (Calcein green) and red (Calcein red) stained PC3 prostate cancer cells. The cells are introduced into the printer as a mixed suspension and encapsulated into droplets randomly. The cell concentrations are set so that most droplets are empty but 5% contain a single red or green cell. They are scanned for fluorescence immediately after generation, to determine if they contain a single cell, appearing as bright red or green droplets (Fig. 2-7a). Empty droplets appear as the dominant feature on the scatterplot at low fluorescence values, while droplets containing red and green cells appear double positive at the upper-right of the scatter plot. In addition to a single cell, each array location is printed with 3 buffer droplets, to bring it to a final volume of ~1 nL. To generate defined arrays of printed cells, we program the instrument with an appropriate file, generating 100 positions containing alternating single red and green cells, a subsection of which is shown in Fig. 2-7b. We confirm proper loading of the array by imaging and find that 98 contain the correct

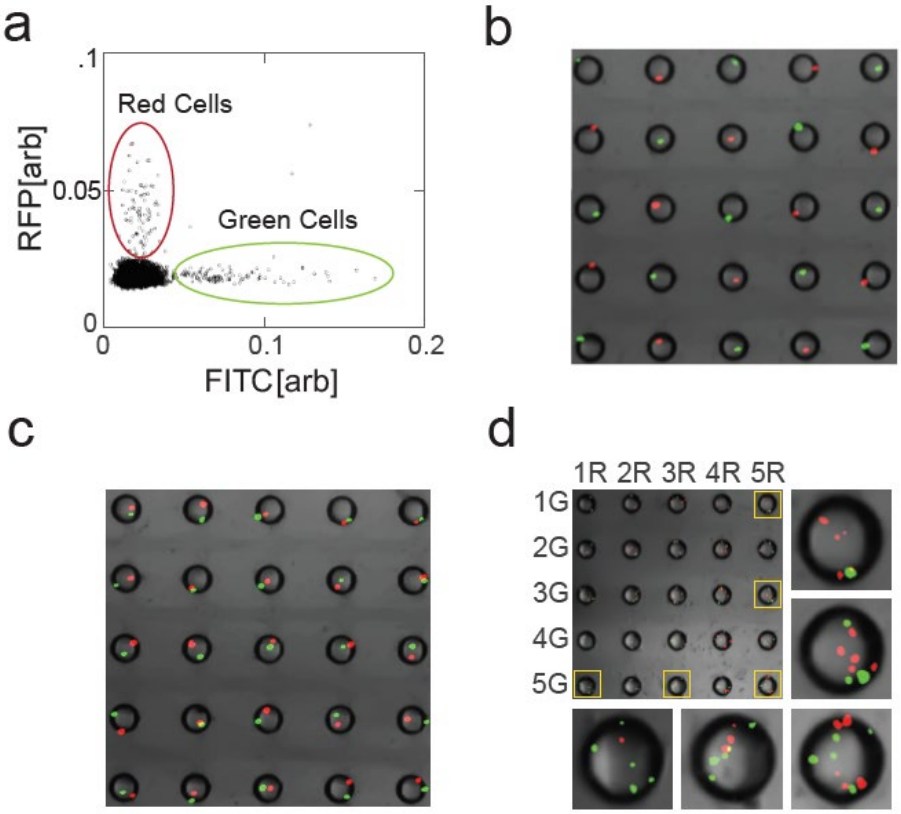

Figure 2-7. Cell printing.
a) Flow cytometry data for an emulsion formed from a mix of calcein green- and calcein red-stained PC3 cells. b) 25-position array where every position is printed with alternating green- and red-stained single cells. c) 25-position array where every well is printed with both a single green- and red-stained cell. d) 25-position array printed with 1-5 green-stained cells on the vertical axis, and 1-5 red-stained cells on the horizontal axis.

contents. Close inspection reveals that positions missing cells have fabrication defects, an artifact of the laser etching method used to pattern electrodes, that should be eliminated when more reliable lithographic wet etching methods are used. By comparison, randomly loading an array with a suspension that contains an average of 0.5 red and 0.5 green cells per each array position, should lead to 36.8% empty wells, 36.8% wells with single cells, and 26.4% of wells with 2 or more cells. In practice, single cells are usually loaded into microscale reactors at ≤ 0.1 cells / well limiting dilution, leading to > 90% empty reactors.

In addition to the study of cell-cell interactions, numerous applications require the deterministic loading of defined numbers of cells or beads into droplets and wells, including new high throughput assays to detect secreted proteins for biologic drug screens (19) or to purify, barcode, and sequence the transcriptomes of thousands of single cells (20). The encapsulation of combinations of discrete entities like cells and beads, however, is usually inefficient, yielding a large majority of improperly loaded droplets. These droplets waste cells and reagents, and can contaminate putatively single cell data with data from multiple cells that otherwise appear indistinguishable.

PDM provides a robust way to deterministically generate any desired combination of cells and beads. The discrete entities are dyed to make them distinguishable and sorted-on-demand as needed to print the combinations, such as a red and green cell at every position (Fig. 2-7c). Of 100 printed positions, we again find 98 contain exactly one cell of each type. Randomly loading a nanowell array with a suspension containing an average of one cell of each type for every well would lead to 13.5% of the wells containing exactly one red and one green cell. Loading wells with two cell types at ≤ 0.1 cells / well limiting dilution, leads to $> 99\%$ of the wells containing undesired contents. This example illustrates the compounded advantage of deterministic cell delivery in constructing multicellular experiments.

Using PDM, arrays containing complicated combinations of cells are no more difficult to construct than single cell arrays and use exactly the same hardware: the print file is simply updated to specify the requisite printing instructions, and the printer runs them. This enables entirely new experiments that depend on multiple cell types. To illustrate the ease with which PDM can accomplish this otherwise impossible feat, we print complex combinations of cells: in which the number of cells per position systematically increases along an axis (Fig. 2-7d). This

demonstrates that a large and reproducible variety of cellular consortia may be produced from a single mixed cell suspension.

2.4 Discussion

A common limitation in flowing droplet microfluidics is that the approach is only compatible with assays that provide bright signals and a high signal-to-noise ratio. Until now, this has restricted the approach to fluorescence or uncommonly sensitive absorbance assays, thus limiting the generality and usefulness of the technique. A critical advantage of PDM is that, as we show, the static droplets can be imaged and observed for long durations, which allows integration of weak signals into the detectable range. Imaging-based methods for absorbance and chemiluminescence have been demonstrated and should be readily applicable to printed droplets (21, 22). Similarly, optical spectrographic methods, like Raman spectroscopy, that are too weak to detect in flowing droplets, have been demonstrated using imaging and should adapt well to printed droplet arrays (23). More ambitiously, recent studies have performed electrospray ionization mass spectrometry on flowing droplets (24). The arrayed nature of printed droplets and ability to engineer the substrate to enable signal enhancement (25), provide unique opportunities for implementing it in an intelligent and useful way, especially in conjunction with laser ionization methods like MALDI, SAMDI, and NIMS that are compatible with the high throughputs required of future single cell biology applications (25–27). The ability to use a broader number of measurement techniques common to the biological sciences should greatly broaden the generality of droplet microfluidics.

We are just beginning the era of high throughput single cell analysis and its impact on our understanding of biology. But the complexity of biology is almost incomprehensible: In addition to characterizing large numbers of single cells, each cell should be characterized as deeply as

possible, especially by combining many independent forms of single cell analysis. However, it is often challenging to integrate multiple analyses together, especially ones that are destructive. Additionally, at the conclusion of a screen, it is often desirable to recover live cells that can be propagated and studied further, or used to seed a new generation of cultures (28). With PDM, such single cell multi-omic measurements become feasible. Cells can grow to high densities in microfluidic droplets, and droplets can be split and printed to independent positions on the substrate. These subsamples provide multiple clonal colonies seeded from single cells that can be subjected to different analyses, including destructive ones. Even more, a pristine subsample can be preserved, providing live sister cells for recovery at the screen's conclusion. Such workflows would be valuable in synthetic biology to evolve cells to produce molecules that can only be detected by destructive analyses, like mass spectrometry. Indeed, the majority of interesting molecules, including pharmaceuticals and chemical building blocks, fall into this category.

We describe a new kind of microfluidics, PDM, that combines the flexibility and programmability of microliter dispensing with the scalability and single cell sensitivity of flowing droplet microfluidics. The core of this approach is a microfluidic sorter coupled to a motorized stage, used as a deterministic cell and droplet dispenser. In addition to allowing complex arrays of droplets and cells to be generated that are impossible by other means, PDM provides many other advantages, making it valuable for applications across subdisciplines of biology. Moreover, with the throughput increases that will result from easily implemented technical improvements, the approach can generate arrays of high complexity quickly and cost-effectively. This will allow new kinds of analyses on single cells and combinations of different cell types, integrating together multiple measurement modalities.

2.5 Materials and Methods

Print head fabrication. The print heads are fabricated from PDMS using standard soft lithography techniques from photo-patterned multi-layer moulds. For print heads that contain integrated drop makers, a 60 μm tall layer of SU-8 photoresist is spun onto a silicon wafer, baked and patterned with a photomask to provide the drop making geometry. Subsequently, an additional 20 μm of SU-8 is spun onto the wafer, baked, and photopatterned to give 80 μm tall geometry for general fluid handling and insertions of the excitation fiber. After this an additional 150 μm of SU-8 is patterned to provide channels for insertion of optical fibers and the piece of tubing used for the print nozzle. After the wafer is developed to remove unpolymerized photoresist, uncured poly(dimethylsiloxane) (PDMS) (10:1 polymer to cross-linker ratio) is poured over the photo-patterned mould. After curing in an oven at 80°C for 80 min, PDMS moulds are extracted from the masters with a scalpel and punched with access ports using a 0.75 mm biopsy core, then plasma bonded to 1" x 3" glass slides. A 1 cm long segment of 135 μm inner diameter, 350 μm outer diameter PVDF capillary tubing (Paradigm Optics, Vancouver, WA) is laterally inserted into the exit channel of the PDMS device and glued in place using 2 part epoxy. To render the channels hydrophobic, they are treated with Aquapel and baked at 80°C for 10 min.

Optical configuration. The tips of the optical fibers used for fluorescence excitation and detection in the experiments are stripped of their insulating layers using a stripping tool and inserted into their respective insertion channels. The excitation fiber has a cladding diameter of 125 μm, an optical core diameter of 105 μm, and a numerical aperture (NA) of 0.22. The detection fiber has a cladding diameter of 225 μm, an optical core diameter 200 μm, and a NA of 0.39. Excitation light is delivered by coupling fibers to 100 mW, continuous-wave lasers with

wavelengths of 405 nm, 473 nm, 532 nm, and 640 nm (CNI Lasers, Changchun, China). The laser light is combined into a single fiber using multimode fiber combiners (OzOptics, Ottawa, Canada). The laser power was adjusted so that that power of each laser color reaching the end of the excitation fiber is ~ 1 mW. Emitted light gathered by the detection fiber is ported to an optical setup where it is columnated, filtered by a quad-bandpass filter to eliminate stray laser light (Semrock, Rochester, NY), and redirected by a series of dichroic filters (Semrock) to detection PMTs (Thorlabs, Newton, NJ). Bandpass filters placed in front of the PMTs give detection channels centered at 448 nm, 510 nm, 571 nm, and 697 nm, roughly corresponding to DAPI, FITC, RFP, and Cy5 channels.

Substrate fabrication. The pattern for laser etched indium tin oxide electrode arrays are designed using Autocad (Autodesk, San Rafael, CA). The array consisted of a 2500 electrode array on a 400 μm grid. 2" x 3" ITO substrates were procured (Delta Technologies, Loveland, CO) and laser patterned by a vendor (Potomac Photonics, Halethorpe, MD). A 10 μm SU-8 layer was spun onto the patterned ITO substrate, baked, and exposed to uniform UV to provide an insulating dielectric layer. Next, a 70 μm tall layer of SU-8 was applied, baked, and patterned to give 150 μm wells centered over the bipolar electrodes locations. Substrates are affixed to petri dishes using kapton tape, and electrical contacts are made with the contact pads using copper tape with a conductive adhesive (Ted Pella, Redding, CA). High voltage AC is applied to the substrate by the signal of a function generator amplified by a 100-1 high voltage amplifier (Model 2220, Trek, Lockport, NY). Typically, the substrate is powered by 200 V - 600 V at 20 kHz.

Printing reagents. The carrier oil used for the printing experiments is Novec HFE-7500 (3M, St Paul, MN) supplemented with 0.2% w/w biocompatible surfactant for bias and spacer oil and with 2% surfactant for dropmaking oil. Before printing the substrate is covered with several millimeters of HFE-7500 with no surfactant to serve as a vapor barrier. Uniform aqueous fluids used for dye printing experiments consist of PBS labeled with 20 μM dextran-conjugated blue dye (CB Cascade blue, Thermo Fisher, Waltham, MA). The dyes used for Fig 3a-c contain an additional 100 μM fluorescein (FITC) or 100 μM dextran-conjugated tetramethylrhodamine (TRITC, Thermo Fisher). The droplet library generated for Fig 2d-f 100 μM FITC, 100 μM TRITC, 100 μM dextran-conjugated Cy5 dye (Alexa 647, Thermo Fisher), or no additional dyes. The large, 10,000 position array in Fig 3g used droplets that contained 25% glycerol in PBS and 100 μM FITC.

Preparation of mixed cell suspensions for single cell printing. Cell printing experiments were performed using PC3 prostate cancer cells cultured in 75 cm^2 flasks in presence of Dulbecco's Modified Eagle Medium (DMEM) containing 4.5 g/L glucose, 0.584 g/L L-glutamine and 3.7 g/L $NaHCO_3$. DMEM was supplemented with 5% fetal bovine serum (FBS) and antibiotics. Cells were grown to 75% confluency prior to Calcein AM staining. Cells were briefly treated with 3 ml 0.25% Trypsin-EDTA to dislodge the cells from culture flask and were washed with 10 ml of DMEM media containing 10% FBS. Cells were resuspended in 5 ml of cold PBS containing 0.2% FBS and cell count and cell viability was measured via TC20 automated cell counter (BioRad, Hercules, CA). One million cells were stained with 10 μM final concentration of either Calcein green, AM or Calcein red-orange, AM (Thermo Fisher) in 100 μl of PBS/10% FBS. Cells were stained for 30 min on ice. Cells were washed two times with 1 ml of cold PBS/10% FBS. Calcein green, AM and Calcein red-orange, AM stained cells were resuspended

and mixed in 200 µl of cold PBS/10% FBS. Calcein, AM staining and cell viability was examined via EVOS cell imaging systems (Thermo Fisher). 200,000 cells were resuspended in 1 mL of PBS that contained 17% Optiprep (Sigma, St. Louis, MO), 1% FBS, and 20 µM CB blue dye. The final cell concentration leads to roughly 1 droplet in 20 containing a single cell.

Preparation of cell suspensions for calcium release assay. PC3 prostate cancer cells were cultured 75 cm^2 flasks in the presence of Dulbecco's Modified Eagle Medium (DMEM) containing 4.5 g/L glucose, 0.584 g/L L-glutamine and 3.7 g/L $NaHCO_3$. DMEM was supplemented with 10% fetal bovine serum (FBS) and antibiotics. PC3 cell were trypsinized using 0.25% (W/V) Trypsin- 0.53 mM EDTA. Cells were collected and resuspended in HBSS with no calcium and no magnesium. 1×10^6 cells were stained with CellTracker Deep Red dye (Thermo Fisher Scientific), final dye concentration 1 µM. Cells were washed three times using 7 mL of cold HBSS with no calcium and no magnesium. 200,000 CellTracker stained cells used for further staining with Fluo-8 No Wash Calcium Assay Kit (Abcam) as instructed in the protocol. The double stained cells were resuspended in cold HBSS with no calcium and no magnesium to final volume of 100 µL. The Fluo-8 stained cells were resuspended in 1 mL of PBS that contained 17% Optiprep (Sigma, St. Louis, MO), 1% FBS, and 20 µM CB blue dye. The final cell concentration leads to roughly 1 droplet in 20 containing a single cell. A second solution containing HBBS, 17% Optiprep, and 250 mM KCl was mixed and used in the experiments to induce cellular calcium release.

Deposition rates for proof of concept dye and cell printing experiments. The rate at which wells may be populated with droplets is currently limited by the speed at which the stage can accurately realign the print nozzle with nanowells on the printing substrate. The stage used in

these studies can address an array with 400 μm spacing at a maximum speed of ~ 4 Hz. The printing experiments performed for Fig 2b, Fig 2c, Fig 2e, and Fig 2f deposited 270 pL droplets at well printing rates of 2-3 Hz. The large array in Fig 2g printed single 500 pL droplets to each well position at a more conservative rate of 1.5 Hz. Because printing is performed in a petri dish open to the environment, fibers or pieces of dust occasionally interfere with printing on small regions of the printing substrate, requiring the printing run to be stopped and the offending particles to be removed. For this reason, the image in Fig. 2g was stitched from printed regions imaged from 3 separate print runs. We do not believe this is a fundamental issue with the technology and will be easily remedied by performing print runs under a hood and by instating robust substrate cleaning procedures. The cell images in Fig. 3 were printed at a rate of ~1 Hz per cell deposited, and was primarily limited by the frequency of cell containing droplets passing through the sorter.

2.6 References

1. Agresti JJ, Antipov E, Abate AR, Ahn K, Rowat AC, Baret JC, et al. Ultrahigh-throughput screening in drop-based microfluidics for directed evolution. Proceedings of the National Academy of Sciences of the United States of America. 2010 Mar 2; 107(9): 4004-4009.

2. Sarkar A, Kolitz S, Lauffenburger DA, Han J. Microfluidic probe for single-cell analysis in adherent tissue culture. Nature Communications. 2014 Mar 5; 5(1): 1-8.

3. Wang BL, Ghaderi A, Zhou H, Agresti J, Weitz DA, Fink GR, et al. Microfluidic high-throughput culturing of single cells for selection based on extracellular metabolite production or consumption. Nature Biotechnology. 2014 Apr 6; 32(5): 473-478.

4. Kim H, Jebrail MJ, Sinha A, Bent ZW, Solberg OD, Williams KP, et al. A Microfluidic DNA Library Preparation Platform for Next-Generation Sequencing. PLoS ONE. 2013 Jul 22; 8(7): e68988.

5. Ng AHC, Dean Chamberlain M, Situ H, Lee V, Wheeler AR. Digital microfluidic immunocytochemistry in single cells. Nature Communications. 2015 Jun 24; 6(1): 1-12.

6. Küster SK, Pabst M, Jefimovs K, Zenobi R, Dittrich PS. High-resolution droplet-based fractionation of nano-LC separations onto microarrays for MALDI-MS analysis. Analytical Chemistry. 2014 May 20; 86(10): 4848-4855.

7. Wang J, Zhou Y, Qiu H, Huang H, Sun C, Xi J, et al. A chip-to-chip nanoliter microfluidic dispenser. Lab on a Chip. 2009 Jul 7; 9(13): 1831-1835.

8. Sciambi A, Abate AR. Accurate microfluidic sorting of droplets at 30 kHz. Lab on a Chip. 2015 Jan 7; 15(1): 47-51.

9. Gu H, Duits MHG, Mugele F. Droplets Formation and Merging in Two-Phase Flow Microfluidics. International Journal of Molecular Sciences. 2011 Apr 15; 12(4): 2572-2597.

10. Abate AR, Hung T, Marya P, Agresti JJ, Weitz DA. High-throughput injection with microfluidics using picoinjectors using picoinjectors. Proceedings of the National Academy of Sciences of the United States of America. 2010 Nov 9; 107(45): 19163-19166.

11. Agrawal S, Cifelli S, Johnstone R, Pechter D, Barbey DA, Lin K, et al. Utilizing Low-Volume Aqueous Acoustic Transfer with the Echo 525 to Enable Miniaturization of qRT-PCR Assay. Journal of laboratory automation. 2016 Feb 1; 21(1): 57-63.

12. Nguyen HQ, Baxter BC, Brower K, Diaz-Botia CA, DeRisi JL, Fordyce PM, et al. Programmable Microfluidic Synthesis of Over One Thousand Uniquely Identifiable Spectral Codes. Advanced Optical Materials. 2017 Feb 2; 5(3): 1600548.

13. Teh SY, Lin R, Hung LH, Lee AP. Droplet microfluidics. 2008 Jan 29..

14. Seidel M, Gauglitz G. Miniaturization and parallelization of fluorescence immunoassays in nanotiter plates. 2003 Jun 1..

15. Du W, Li L, Nichols KP, Ismagilov RF. SlipChip. Lab on a Chip. 2009 Aug 21; 9(16): 2286-2292.

16. Ochs CJ, Abate AR. Rapid modulation of droplet composition with pincer microvalves. Lab on a Chip. 2015 Jan 7; 15(1): 52-56.

17. Brouzes E, Medkova M, Savenelli N, Marran D, Twardowski M, Hutchison JB, et al. Droplet microfluidic technology for single-cell high-throughput screening. Proceedings of the National Academy of Sciences of the United States of America. 2009 Aug 25; 106(34): 14195-14200.

18. Gracz AD, Williamson IA, Roche KC, Johnston MJ, Wang F, Wang Y, et al. A high-throughput platform for stem cell niche co-cultures and downstream gene expression analysis. Nature Cell Biology. 2015 Mar 2; 17(3): 340-349.

19. Konry T, Golberg A, Yarmush M. Live single cell functional phenotyping in droplet nano-liter reactors. Scientific Reports. 2013 Nov 11; 3(1): 1-5.

20. Klein AM, Mazutis L, Akartuna I, Tallapragada N, Veres A, Li V, et al. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. Cell. 2015 May 30; 161(5): 1187-1201.

21. Lee JY, Clarke ML, Tokumasu F, Lesoine JF, Allen DW, Chang R, et al. Absorption-based hyperspectral imaging and analysis of single erythrocytes. IEEE Journal on Selected Topics in Quantum Electronics. 2012; 18(3): 1130-1139.

22. Li H, He Z. Magnetic bead-based DNA hybridization assay with chemiluminescence and chemiluminescent imaging detection. Analyst. 2009 Mar 23; 134(4): 800-804.

23. Duncan MD, Reintjes J, Manuccia TJ. Imaging Biological Compounds Using The Coherent Anti-Stokes Raman Scattering Microscope. Optical Engineering. 1985 Apr 1; 24(2): 242352.

24. Sun S, Slaney TR, Kennedy RT. Label free screening of enzyme inhibitors at femtomole scale using segmented flow electrospray ionization mass spectrometry. Analytical Chemistry. 2012 Jul 3; 84(13): 5794-5800.

25. Heinemann J, Deng K, Shih SCC, Gao J, Adams PD, Singh AK, et al. On-chip integration of droplet microfluidics and nanostructure-initiator mass spectrometry for enzyme screening. Lab on a Chip. 2017 Jan 21; 17(2): 323-331.

26. Amantonico A, Urban PL, Fagerer SR, Balabin RM, Zenobi R. Single-cell MALDI-MS as an analytical tool for studying intrapopulation metabolic heterogeneity of unicellular organisms. Analytical Chemistry. 2010 Sep 1; 82(17): 7394-7400.

27. Patel K, Sherrill J, Mrksich M, Scholle MD. Discovery of SIRT3 Inhibitors Using SAMDI Mass Spectrometry. Journal of biomolecular screening. 2015 Aug 25; 20(7): 842-8.

28. Nyyssönen M, Tran HM, Karaoz U, Weihe C, Hadi MZ, Martiny JBH, et al. Coupled high-throughput functional screening and next generation sequencing for identification of plant polymer decomposing enzymes in metagenomic libraries. Frontiers in Microbiology. 2013 Sep 23; 4(SEP): 282.

Chapter 3 : Linked optical and gene expression profiling of single-cells at high-throughput.

Elements of this chapter have been reprinted from the manuscript “Linked optical and gene expression profiling of single-cells at high-throughput,” by Jesse Q. Zhang, Christian A. Siltanen, Leqian Liu, Kai-Chun Chang, Zev J. Gartner, and Adam R. Abate, published in *Genome Biology*, volume 21, article number 49 on February 24, 2020.

3.1 Abstract

Single-cell RNA-sequencing has emerged as a powerful tool for characterizing cells, but not all phenotypes of interest can be observed through changes in gene expression. Linking sequencing with optical analysis has provided insight into the molecular basis of cellular function, but current approaches have limited throughput. Here, we present a high-throughput platform for linked optical and gene expression profiling of single cells. We demonstrate accurate fluorescence and gene expression measurements on thousands of cells in a single experiment. We use the platform to characterize DNA and RNA changes through the cell cycle and correlate antibody fluorescence with gene expression. The platform’s ability to isolate rare cell subsets and perform multiple measurements, including fluorescence and sequencing-based analysis, holds potential for scalable multi-modal single cell analysis.

3.2 Introduction

Cellular processes such as replication, migration, and differentiation, are tightly controlled by signaling and gene regulatory networks (1,2,3). These processes are dynamic, and at any point a cell may exist along a continuum of states (4). Thus, cell state heterogeneity is often masked when bulk methods are used to analyze populations (5,6). The development of high

throughput single cell RNA sequencing (scRNA-seq) has enabled populations to be analyzed at the single cell level (7,8,9), facilitating the dissection of cellular heterogeneity and the construction of an atlas of cell states across the human body (10). However, gene expression is just one dimension by which cells may be characterized, and many properties, such as epigenetic state, protein expression, enzyme activity, and cellular morphology, are not readily measured by scRNA-seq (11,12).

More comprehensive cell characterization can be accomplished by combining scRNA-seq with complimentary measurement modalities. Optical approaches, including microscopy and flow cytometry, can characterize morphological and fluorescence phenotypes prior to scRNA-seq (13,14). Linked optical analysis and scRNA-seq has been applied to *in vitro* cultures, patient tissues, and stem cells, revealing molecular links to cellular function (15,16,17). While powerful, these approaches are limited in throughput (15,18). Cytometry methods are more scalable since the instrument can automatically sort cells into wells for automated library preparation, increasing throughput to hundreds of cells (16,17,19). Going beyond this, however, is impractical because the time and volume of reagent required to process tens or hundreds of thousands of cells is prohibitive (20). Recent spatial transcriptome sequencing approaches might ultimately enable scalable imaging and scRNA-seq but rely on methods to image and label cells that are not yet standard in the field (21,22).

Microwell technologies improve throughput and can process thousands of single cells for RNA sequencing with the benefit of allowing cells to be imaged prior to sequencing preparation (23,24). However, because these platforms load cells stochastically, most wells remain empty, limiting the total number of cells analyzed to ~2,000. Moreover, there is no integrated enrichment in these methods, making it challenging to isolate rare cell subsets, which are

important in a variety of applications, including cancer pathophysiology, immunology, and stem cell biology (25,26,27). For example, for a target cell present at 5%, only ~100 cells would be captured, yielding a throughput no better than flow cytometric methods. To enable optical and sequence-based analysis of specific cells in heterogeneous samples, a new approach is needed that can specifically isolate, perform optical measurements on, and sequence large numbers of target cells.

In this paper, we present high-throughput optical and RNA-sequencing analysis of single cells. Our instrument functions like a flow cytometer, optically scanning cells in flow and sorting targets into individual nanoliter wells for sequencing preparation. To pair the optical and sequencing data, the wells are indexed with coordinate oligos that are captured during sequencing; this allows all reads for a given cell to be associated with a specific well on the plate, thereby pairing it with the optical data. The cell analysis is accomplished at ~1 KHz and dispensing at ~5 Hz, allowing hundreds of rare cells to be isolated in a few minutes. The total volume of reagent on the array is ~1 μL per 1,000 single cell transcriptomes, representing a 1000-fold reduction compared to conventional microliter well plates. Moreover, with standard fabrication techniques, ~10,000 wells can be fabricated on a microscope slide, providing a scalable means by which to acquire linked single cell optical and gene expression data for selected cell populations.

3.3 Results and Discussion

3.3.1 Using Printed Droplet Microfluidics (PDM) for linking single-cell optical and gene expression measurements

Our single cell analysis platform is based on Printed Droplet Microfluidics (PDM) (28,29), an approach that allows cells to be optically scanned and dispensed to custom nanoliter

well plates (nanoplates) (**Fig. 3-1A**). To perform linked optical and scRNA-seq analysis, we record the fluorescence of a cell while confined to a droplet, then dispense the cell and droplet to the nanoplate at defined locations. Then, scRNA-seq library preparation is performed on each cell using specific "coordinate oligos" encoding each cell's location on the nanoplate. After sequencing, these oligos allow each cell barcode to be traced to a well of origin, thereby linking it to the optical data collected for that cell. The workflow is similar to flow cytometry, except that the sorter is a microfluidic device and the wells in which the cells are dispensed are ~10,000-fold smaller than conventional microwells. This reduction in volume, combined with the speed of the microfluidic printer, enable highly scalable optical phenotyping and sequencing of single cells.

Prior to device operation, a separate flow focusing device encapsulates cells in a droplet emulsion. We introduce this emulsion into the PDM device, where each drop is optically scanned (**Fig. 3-1A, left**). As in flow cytometry, laser-induced fluorescence accomplishes the optical analysis, whereby focused lasers excite fluorescence of the cells which a multicolor detector then captures (**Fig. 3-1A, middle**). Cells with desired fluorescence properties are isolated through sorting them into a printing nozzle that dispenses them into a nanowell on the substrate (**Fig. 3-1A, right**). We record the cell fluorescence and dispense location, allowing this information to be paired with the scRNA-seq data collected later.

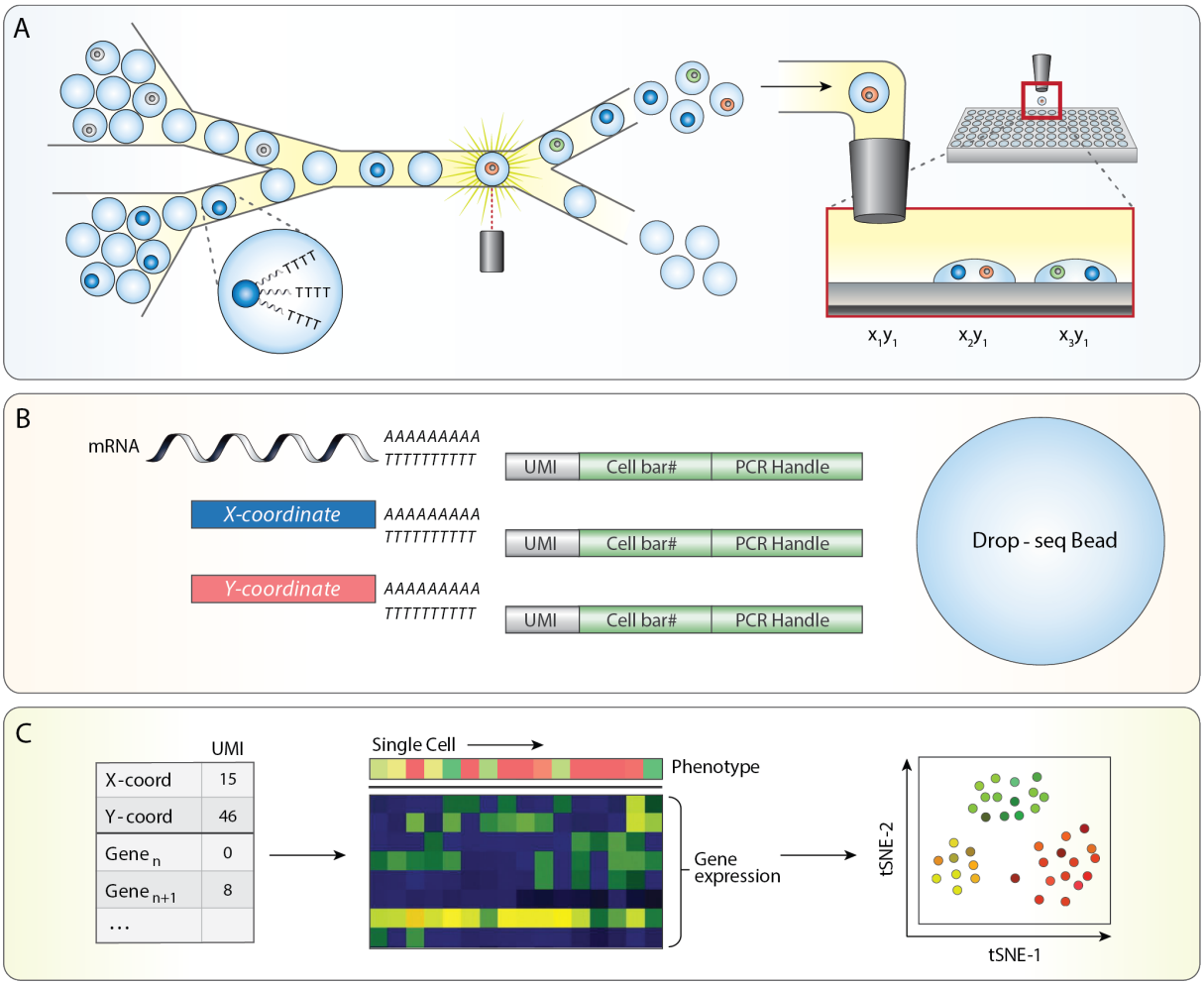

Figure 3-1. A high throughput platform for linked optical phenotype and gene expression of single cells. (A) Monodisperse droplet emulsions containing encapsulated poly-T mRNA capture beads and cells are input into a microfluidic device. Fluorescence signal from droplets is interrogated and used to selectively dispense a cell and bead to indexed locations on a nanowell array. (B) Each bead binds mRNA from cell lysate as well as a unique combination of poly-A barcode oligos denoted by nanowell coordinate. (C) UMI counts on each bead are collected through sequencing into an expression matrix for each cell. Nanowell coordinate is assigned based on the abundance of barcode oligos and paired with fluorescence data obtained during cell sorting, which enables downstream linked analyses such as dimensionality reduction visualizations of gene expression paired with optical phenotype.

To link the optical and sequencing data, we index the wells such that each cells' dataset can be traced back to a well on the array. The indexes comprise "coordinate oligos" pre-loaded into the wells using a commercial reagent spotter (**Fig. 3-1B**, **lower left**) (30). To index the array, we place "X" and "Y" coordinate oligos, each of which contains a different 8 base sequence encoding the specific location of a given nanowell on the plate. The coordinate oligos are polyadenylated, allowing them to be captured along with cellular mRNA during the scRNA-seq

library preparation. For scRNA-seq, we adapt the validated "Drop-Seq" protocol (7), which uses beads coated with poly-deoxythymidine "barcode" oligos to capture and label both mRNA and coordinate oligos. We accomplish this by co-dispensing beads and cells in each nanowell and lysing the cells. After retrieving the beads, performing the requisite library preparation steps of Drop-Seq, and sequencing the barcoded cDNA, we obtain a collection of reads representing the cell transcriptome and coordinate oligos, all sharing a Drop-Seq barcode. Thus, the location of the cell from which the data originate is encoded in the sequencing data, allowing it to be traced back to a specific well on the array (**Fig. 3-1C, left**) and associated with the previously recorded optical data (**Fig. 3-1C, middle**). With this paired dataset, we can use dimensionality reduction methods to first visualize gene expression data, to which we add the optical phenotype information (**Fig. 3-1C, right**).

3.3.2. PDM device operation for dispensing of beads and cells to nanowell array.

The microfluidic print device consists of a droplet spacer, sorter, and printing nozzle (**Fig. 3-2A**). A packed emulsion containing cells or beads is introduced, spaced by oil, and optically scanned by a four-color laser-induced fluorescence detector (**Fig. 3-2A, red outline**). Embedded fiber optics excite and collect fluorescence that is processed through filters and analyzed in real time by custom software; this allows cell, bead, and droplet fluorescence and scattering to be recorded, to determine whether to print the droplet and its contents to the current nanowell. Printing is achieved by sorting a droplet (**Fig. 3-2A, green outline**) into the printing nozzle positioned above a nanowell (**Fig. 3-2A, purple outline**); if the current droplet should not be printed, it is not sorted into the nozzle and passes into the discard channel. Because the carrier oil is viscous and denser than water, in the absence of other forces, the ejected droplet would

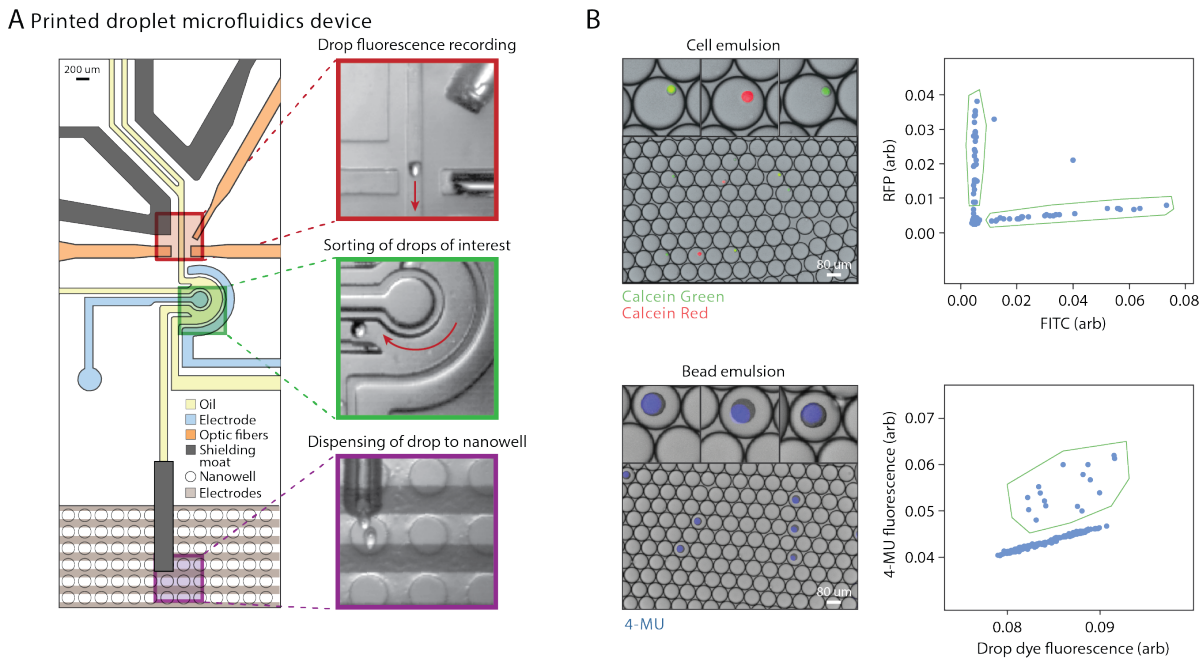

Figure 3-2. Printed Droplet Microfluidics (PDM) operation for deterministic loading of nanowell array with beads and cells.
(A) An inset of the microfluidic device aligned over the nanowell array, with images (top to bottom) of regions of drop fluorescence recording, sorting of drops of interest, and dispensing of drops to nanowells. (B) Monodisperse droplet emulsions containing fluorescently labeled cells (top) or beads (bottom) are input into PDM. Drops of interest (insets) are enriched for by gating on fluorescence plots (right) generated during device operation.

float away and not go into the nanowell. Thus, to dispense it into the nanowell, electrodes positioned under the substrate emit an oscillating electric field. This field pulls the dispensed droplet into the nanowell by a dielectrophoretic force and is key to the speed of PDM, since it allows a droplet to travel the final few hundred microns from the printing nozzle to the nanowell in tens of milliseconds (28). Moreover, because the trap extends above the substrate, the printer need not dispense the droplets with perfect accuracy into the nanowells, since any droplet within the electric field will, ultimately, be pulled into the nearest nanowell. The trapping field also ensures that the printed droplets remain fixed in the wells. Upon completion of a print run, droplets can be released by un-powering the electrodes.

To demonstrate the accuracy of scRNA-seq using our approach, we perform an experiment with two cell-types. We prepare and encapsulate a mixed suspension of Calcein Red

stained mouse (3T3) and Calcein Green stained human (HEK293) cells (**Fig. 3-2B, upper left**). When scanned in the print head, we observe distinct green and red cell populations (**Fig. 3-2B, upper right**); thus, with suitable gating instructions, the printer can print these cells in a defined pattern to the nanoplate. To enable scRNA-seq of the printed cells, Drop-Seq beads must also be printed, which requires that they be detectable in the print head; this is accomplished by labeling them with 4-MU, a blue dye that does not overlap with the cell stains (**Fig. 3-2B, lower left**). The 4-MU dye is insoluble in water and remains sequestered within the beads following labeling, allowing bead-loaded to be discerned from bead-empty droplets (**Fig. 3-2B, lower right**). To print cells and beads in defined combinations, we generate a "print file" containing gating and location instructions that we input into the printer software; the printer reads this file, printing cells and beads to the nanoplate according to the instructions in the file.

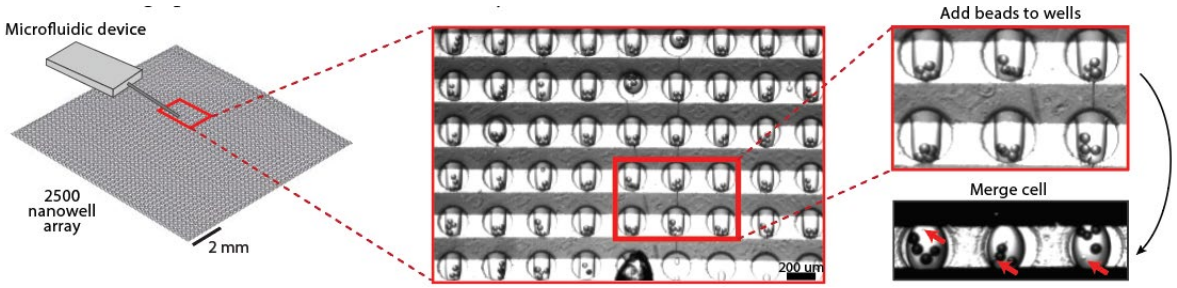

Figure 3-3. Loading of nanowell array with cells and beads.
First beads are added to nanowells, followed by merging of a cell-containing drop in lysis buffer. To improve mRNA capture efficiencies, multiple beads can be controllably printed into each nanowell.

3.3.3. Improving cell capture efficiency by printing multiple beads per well.

A strength of PDM for scRNA-seq is that it allows the nanowells to be controllably loaded with cells and beads; this contrasts with other high throughput scRNA-seq methods which randomly load cells or beads and, thus, are less efficient. Moreover, PDM allows systematic variation of nanowell contents across the array, to choose conditions that maximize data quality (28). For instance, minimizing the rate of doublets – when two cells are inadvertently sequenced as one – leads to better data quality and lower per-cell sequencing costs (31,32). While we

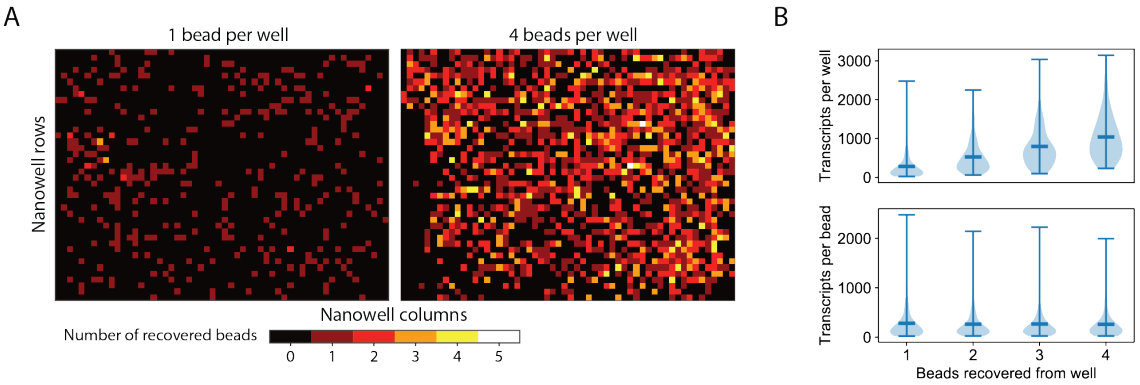

Figure 3-4. Improved cell and mRNA capture when printing multiple beads to each well.
(A) One or four beads were printed to each well of a 42 by 56 nanowell array along with an alternating pattern of mouse and human cells. The number of recovered beads per nanowell position was determined by the number of unique cell barcodes mapped back to each nanowell. (B) When printing four beads per well, the distribution of transcripts recovered from each nanowell were calculated as a function of the number of beads recovered. The distribution of the number of transcripts originating from each bead within a nanowell was also plotted as a function of the number of beads recovered per nanowell.

cannot directly monitor the number of cells passing through each droplet, in the two-cell experiment we gate out heterotypic doublets of droplets containing a red and green cell and minimize the rate of homotypic doublets by excluding the upper tail of the Calcein Red and Calcein Green distributions. Moreover, controlled printing allows us to load multiple capture beads to every well (**Fig. 3-3**), which can increase cell and mRNA capture efficiency by compensating for losses during sequencing library preparation (33). To illustrate this, we print two substrates, the first with one bead per well and the second with four, both on 42 by 56 nanowell (2352) arrays. All beads originating from the same well contain the same coordinate barcodes, allowing us to group reads associated with multiple beads together. Due to loss of beads during library preparation, starting with more beads per well increases the likelihood of recovering at least one bead from every well (**Fig. 3-4A**). When printing more beads, we also recover more transcripts per well (**Fig. 3-4B, upper**), and that the number of transcripts per bead remains consistent when recovering up to 4 beads per well (**Fig. 3-4B, lower**). This suggests that increasing bead surface area per cell lysate increases mRNA capture.

3.3.4. Accuracy of single-cell sequencing and optical phenotype linking confirmed with two-cell experiment.

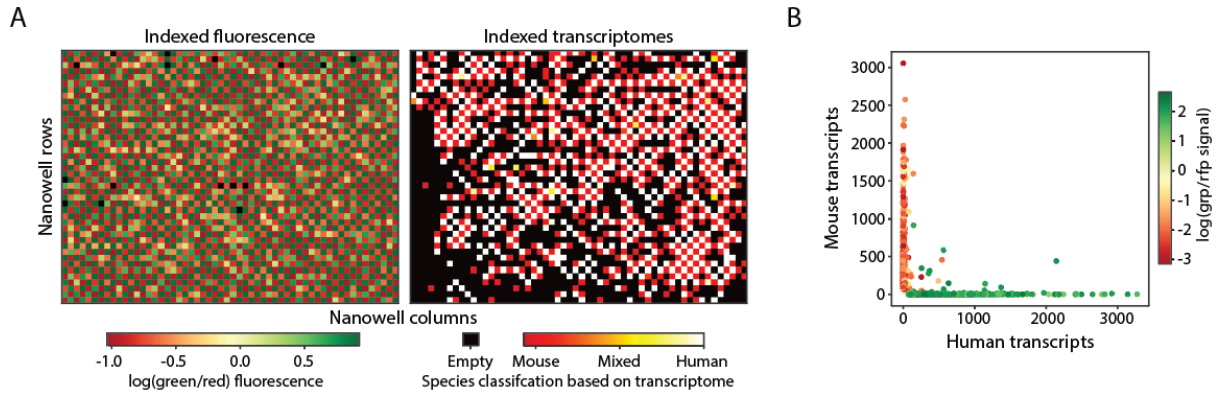

Figure 3-5. Linked optical phenotype and gene expression measurements verified with two species experiment.
(A) Left: Fluorescence data from alternating printing of Calcein green stained human cells and Calcein red stained mouse cells indexed by nanowell position. Right: Ratio of human to mouse transcripts recovered from each nanowell based on printing four beads per nanowell. (B) Transcript counts by nanowell position are annotated with the green-red fluorescence ratio from the cell printed into the corresponding nanowell.

Single cell RNA-seq methods depend on minimal cross contamination of RNA between cells (31). To investigate cross contamination, we print mouse and human cells in a checkerboard pattern and classify each transcriptome according to which species' genome the transcripts predominately align. We find that the optical data align 99.3% of the time with the expected printing pattern (**Fig. 3-5A, left**). We recover transcripts from 1204 nanowells following sequencing and bioinformatic analysis, with 97.6% of transcriptomes having species purities of at least 90% mouse or 90% human and matching the expected printing pattern (**Fig. 3-5A, right**). As we demonstrate a near-perfect success rate when printing beads and cells, we reason that the most of the lost nanowells represent beads that were not collected after printing or lost during library preparation. We find that 1.8% of nanowells have less than 90% species purity, suggesting either mRNA cross contamination or misprinting of cells (dots not aligned with axes) (**Fig. 3-5B**). When we sequence deeper on a smaller subset of beads, we recover on average 17500 and 15200 transcripts from mouse and human cells, respectively (**Fig. S3-1**). In sum, we

demonstrate quality optical and scRNA-seq data from over a thousand cells when printing four beads per well.

3.3.5. Investigating cell-cycle associated changes in DNA content and gene expression in cycling Jurkat cells.

Cells undergo changes in state and phenotype through the cell cycle (17). For example, by the G2 phase, cells have doubled their genome, allowing it to be optically detected via DNA staining (34). Moreover, different genes peak and diminish in expression through the cycle, making the cell cycle useful for validating our approach. As a model system, we use Jurkat cells stained with DRAQ5, which is a live-cell stain for genomic DNA (**Fig. 3-6A, upper**). We observe a broad distribution of DNA fluorescence, the brightest of which likely correspond to cells with the most DNA and, thus, in the G2- or M-phase of the cell cycle. To confirm these results, we perform flow-cytometry of the suspension and obtain a similar distribution (**Fig. 3-6A, lower**). We utilize PDM to generate a 56 by 56 nanowell array to which we print a checkerboard pattern of low and high DRAQ5 expressing cells (**Fig. 3-6B**). We sequence transcriptomes from 437 cells and use Uniform Manifold Approximation and Projection (UMAP) to visualize these cells (35). Through assigning cells to G1, S, or G2M phases based on their expression of cell cycle-associated genes and using those genes for principal component analysis, we generate a UMAP plot which identifies three clusters in agreement with three stages of the cell cycle (**Fig. 3-6C, left**). To determine whether these classifications agree with the optical data, we annotate the points of the UMAP plot according to the magnitude of DRAQ5 fluorescence (**Fig. 3-6C, right**). The plots are in general agreement, with the state comprising the

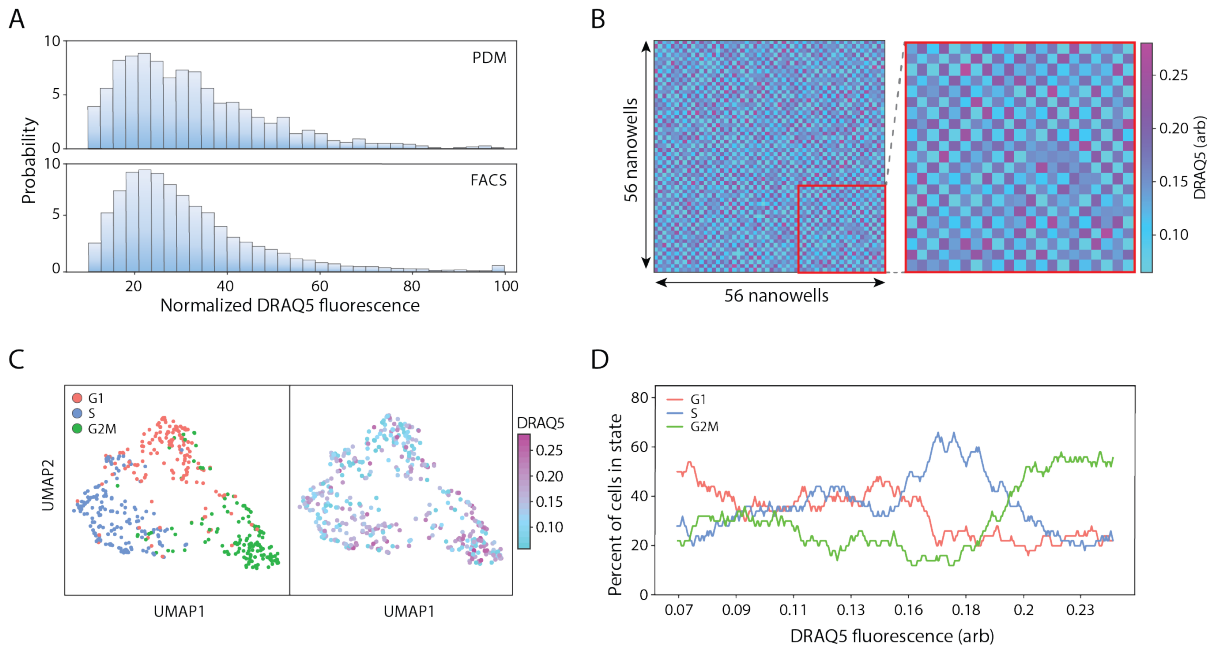

Figure 3-6. Linked fluorescence and gene expression analysis of cell cycle state in Jurkat cells stained with a DNA-binding dye.
(A) The frequency distribution of Jurkat cells stained with DRAQ5 encapsulated within droplets was analyzed on both PDM and a flow cytometer. (B) An alternating pattern of high and low expressing DRAQ5 Jurkats was dispensed to a 56 by 56 nanowell array using PDM. Fluorescence measurements were indexed by nanowell position (inset). (C) Transcriptomes from 498 cells were recovered and clustered based on cell cycle state (left) predicted based on a set of cell cycle dependent genes. DRAQ5 fluorescence data collected during printing was then overlaid (right). (D) Cells were ordered by low to high DRAQ5 signal (top bar), and the fraction of cells in each cell cycle state was calculated over a 50-cell sliding window using corresponding cell cycle state assignments by gene expression analysis.

most DNA (G2M) appearing brightest in the DNA stain. To observe how the population varies through this cycle, we order the cells by fluorescence and plot the proportion in the three states as classified by gene expression (**Fig. 3-6D**). We expect the proportion of cells in G1 to be at low DRAQ5 signal, S phase in the middle of the distribution, and G2 at the top end. The peak of the G1 phase curve is at the low end of the DRAQ5 distribution, S-phase in the middle, and G2/M at the top end. We observe general concordance with the expected trend when we pair fluorescence and scRNA-seq measurements of cell cycle state. With our platform we thus demonstrate characterization of a fundamental biological process through linked optical and gene expression analysis.

3.3.6. Enrichment and paired antibody fluorescence and gene expression analysis of CD14/CD16+ peripheral blood mononucleated cells (PBMCs).

Cell populations are often heterogeneous, with particularly important cells present below 5%, such as those involved in cancer pathogenesis, immunological memory, and maintaining adult stem cell niches (25,26,27). With PDM, we can enrich for rare cell types using their optical phenotype prior to their loading on the nanowell array, markedly increasing the abundance of rare cell types in the obtained sequencing data. We demonstrate this capability by enriching for CD14-positive (CD14+) and CD16-positive (CD16+) peripheral blood mononuclear cells (PBMCs). We count CD14+ and CD16+ cell abundances prior to analysis using PDM and determine that they constitute ~25% and ~3% of all PBMCs, respectively. We load the nanowell array with CD14+ and CD16+ PBMCs, demonstrating an up to 405-fold enrichment of cells over their detection frequencies (**Fig. 3-7A, upper**). We recover transcriptomes from 310 cells and determine their phenotype through gene expression analysis. By overlaying the cell phenotype over the fluorescence data, we determine that natural killer (NK) cells largely are CD14-/CD16+, nonclassical monocytes CD14+/CD16+, and classical monocytes and monocytes of an intermediate phenotype CD14+/CD16- (**Fig. 3-7A, lower**). These four phenotypes cluster separately on a UMAP (**Fig. 3-7B, left**). By contrast, when we sample cells from the sequencing data such that the ratios of CD14+ to CD16+ cells match the observed abundances, we do not resolve the NK or nonclassical monocyte clusters after clustering (**Fig. 3-7B, right**). The ability of our approach to enrich for specific cells in a heterogeneous population and dedicate all the sequencing to them thus provides a powerful advantage when the cells of interest are rare.

Flow cytometry is routinely used to profile monocyte phenotypes based on CD14 and CD16 surface marker expression (36). By overlaying the optical data on the UMAP, we show that CD14+/CD16- cells largely correlate with the classical and intermediate monocyte clusters, and the NK cells and nonclassical monocyte clusters are CD16+ (**Fig. 3-7C**). We also observe an intermediate monocyte population that is CD14+/CD16+. To further study the trends between optical phenotype and gene expression, we plot gene expression versus CD14 or CD16 fluorescence for three relevant genes: LYZ, S100A9, and IFITM2 (**Fig. 3-7D**). LYZ and S100A9 are associated with the classical monocyte phenotype (37), and IFITM2 is associated with nonclassical monocytes (38). By ordering cells from low to high fluorescence and taking the moving average of gene expression, we uncover trends between fluorescence and gene

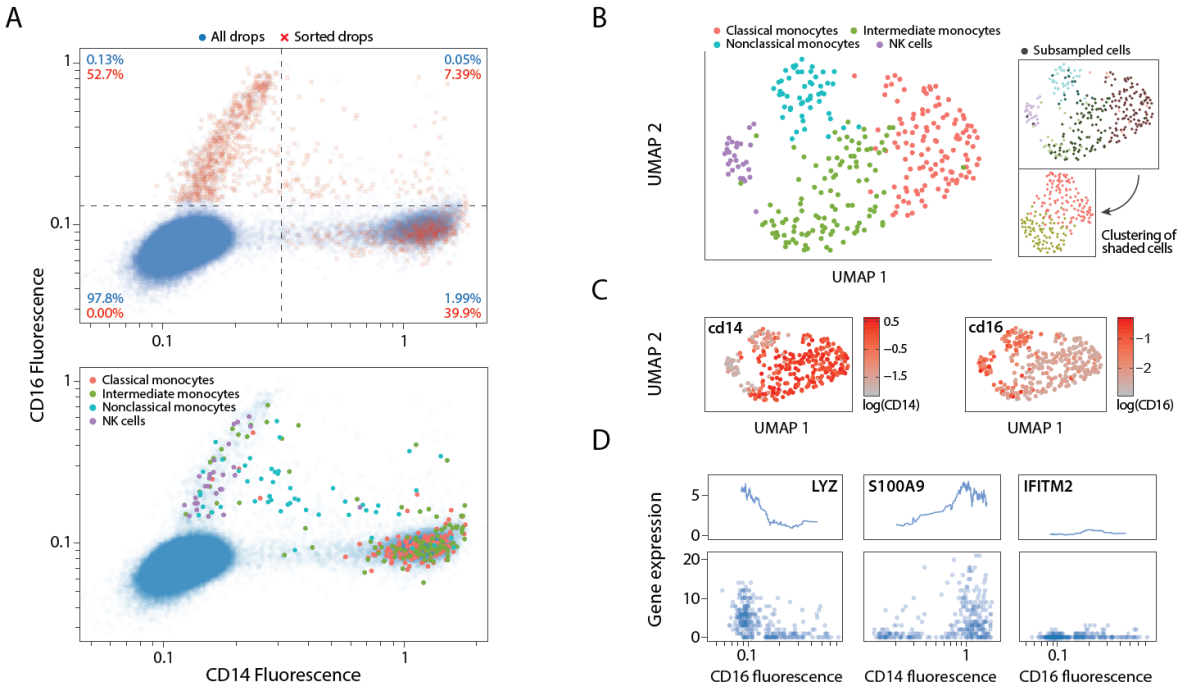

Figure 3-7. Paired optical phenotype and gene expression analysis of enriched CD14+/CD16+ cells from PBMCs.
(A) Use of dual antibody panel and PDM to significantly enrich for CD14+ and CD16+ cells from PBMCs encapsulated within droplets (top). Percentages represent quadrant proportions, with red indicating post-sort frequency. Annotation of sorted cells by cell type as identified through RNA-seq (bottom). (B) UMAP of bioinformatically filtered cells clustered by cell type (left). Cells are subsampled such that the ratio of CD14+ to CD16+ matches the observed pre-sorting ratio; those cells are clustered (right). (C) Overlays of CD14 and CD16 fluorescence for each cell on the UMAP. (D) Scatter plots of gene expression and fluorescence data for 3 selected genes, with moving average of expression plotted above (bin size=30).

expression for these genes. LYZ expression decreases with increasing CD16 fluorescence; S100A9 expression increases with increasing CD14 fluorescence; IFITM2 expression is globally lower than the others, but expression peaks in the middle of the CD16 fluorescence distribution. By combining fluorescence and gene expression measurements, we highlight that the expression of signature monocyte phenotype associated genes gradually changes with optical phenotype.

3.3.7. Conclusion

Single cell RNA-seq is a powerful and general method for analyzing cells, but not all traits of interest are observable in gene expression data alone. Here, we demonstrate an approach that allows gene expression to be linked to optical data in a high-throughput format scalable to thousands of single cells. A key concern for any scRNA-seq workflow that relies on barcoding cells for bulk sample amplification, is the loss of cells due to loss of beads during downstream processing (33). By spreading each cell's transcriptome over several beads, we increase the probability of recovering transcriptomes for all cells. Our platform's ability to localize combinations of cells, beads, and reagents at defined positions on a nanoliter array affords other powerful capabilities, such as systematic variation of cell, reagent, and drug combinations and tuning of optical and sequencing parameters to achieve optimal data. The open nature of the array also makes it amenable to additional measurement modalities, such as atomic force microscopy, mass spectrometry, and chemical assays, all of which can be linked with optical and scRNA-seq data using the approach we have presented (39,40). The ability to link sequencing readouts with other measurements for thousands of single cells will facilitate further investigations into the molecular underpinnings of cell function (41).

3.4 Materials and Methods

Microfluidic device fabrication. PDM chips are fabricated by poly(dimethylsiloxane) (PDMS) molding over a SU-8 master. Briefly, a three-layer SU-8 negative master is patterned to form 20, 80, and 150 um tall features using previously described multi-layer SU-8 photolithography techniques (28). Following casting of PDMS over the SU-8 and curing at 65 degrees for 2 hours, inlet holes are punched into devices using a 0.75 mm biopsy core. Devices are then plasma bonded to 25 mm x 75 mm glass slides. 1 cm of PE/5 tubing (Scientific Commodities) is inserted into the nozzle channel and sealed with a 1-minute instant mix epoxy (Norland). Channels are then treated with AquaPel (AquaPel). Drop-making devices are fabricated as previously described (28). Two devices with a T-junction cross-section of 80 um x 45 um and 80 um x 80 um are used.

Nanoplate fabrication. A negative of the electrode pattern is fabricated on a 50 mm x 75 mm glass slide by positive resist photolithography. A 2 um thick layer of MA-P 1215 (Micro Resist Technology) is spin-coated onto the slide and baked for 1 minute on a 95 degree hotplate. The slide is then exposed to collimated 190 mW UV light (Thorlabs) for 3.5 minutes. The slide is developed in MF-24A developer (Dow Chemical) for 1 minute. Patterned slides then have a 200 A thick layer of chromium deposited on them (LGA Thin Films). The removal of the photoresist with acetone yields the electrode pattern. Nanowells are fabricated on electrode slide by first masking off the regions of electrode contact and spin-coating a 15 um thick layer of uncured PDMS. PDMS is then cured for 3 minutes on a 95 degree hotplate. Following plasma treatment of the slide, a 40 um thick layer of SU-8 is spin coated onto the slide and allowed to soft-bake on a 95 degrees for 10 minutes. The slide is exposed to UV light under a photomask for 90 seconds, followed by 5 minutes of post-exposure baking at 95 degrees. The slide is then immersed in

PGMEA developer (Sigma) for 5 minutes, rinsed with PGMEA and isopropanol, then dried on the hot plate for 2 minutes. Slides are then plasma treated and placed in a petri dish adjacent to reservoirs of trichloro(1H,1H,2H,2H-perfluorooctyl)silane (Sigma) for 2 hours under vacuum at room temperature.

Nanowell coordinate indexing. Nanowells are barcoded using the sciFLEXARRAYER S3 (Scienion AG). A 96 well 'source-plate' containing up to 44 coordinate oligos (**Table S3-1, Table S3-1**) diluted to 1 nM in DI water is prepared. 2 nL of each barcode oligo solution is dispensed to nanowells according to a pre-programmed print routine to label each nanowell with a unique but known combination of three oligos (**Fig. S3-2**). Nanowells are split into 14-by-14 subarrays, of which each subarray had 14 unique x and 14 y coordinate oligos. Subarrays are tiled together, with each subarray having a unique z coordinate oligo, until the array reached the desired size. Following printing, slides are placed in a petri dish and sealed with parafilm and stored at -20 degrees until ready to use.

PDM operation and optical configuration. A multimode excitation fiber with a core diameter of 105 um and a NA of 0.22 (Thorlabs) is inserted into a guide channel in the PDM device. Similarly, an emission detection fiber with core diameter of 200 um and NA of 0.39 (Thorlabs) is inserted into a second guide channel in the PDM device. Four 50 mW continuous wave lasers with wavelengths of 405, 473, 532, and 640 nm are combined and coupled to the excitation fiber. Emitted light is columnated and ported into a quad-bandpass filter, then passed through a series of dichroic mirrors. Bandpass filters of 448, 510, 571, and 697 nm past each dichroic mirror enable wavelength-specific detection of emitted light by PMTs. Electrode channels and a 'Faraday moat' are filled with a 5M NaCl solution. A positive electrode is connected to a function generator and a high voltage amplifier while a second electrode is grounded. Fluidic inputs into

the PDM device are driven by syringe pumps (New Era). Bias and spacer oil containing 0.2% w/w IK in HFE-7500 are flowed through the device at a flow rate of 2000 uL/h. A waste channel is driven with a negative flow rate of -3000 uL/h. Monodisperse droplet emulsions are reinjected into the device at a flow rate of 100 +/- 50 uL/h. Real-time optical signal acquisition through a field programmable gate array (National Instruments) is displayed on a LabView software. Optical signal is processed in real-time and displayed on a fluorescence dot plot, in which drop types of interest can be assigned by specifying gates. Droplets are subsequently sorted by passing a high frequency pulse through a high voltage amplifier (Trek 690E-6). Typical droplet sorting parameters range from 10-20 kHz, 50-100 cycles, and 0.5-1.0 kV. Copper tape with a conductive adhesive (Ted Pella) is affixed to two electrode contact pads on the nanoplate. One pad is connected to ground, while the other one is connected to a function generator and a high voltage amplifier, providing power at 200 - 600V at 20 - 30 kHz. Slides are immersed in a bath of 2% w/w IK in FC-70 (3M) during printing operation.

Cell culture. HEK and 3T3 cells (ATCC) are cultured in 75 square-cm flasks in the presence of Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and 1x Penicillin-Streptomycin at 37 deg and 5% CO2. Cells are treated with 0.25% Trypsin-EDTA and washed with media to generate cell suspensions. The viability and cell concentration are counted by a TC20 automated cell counter (BioRad). Cell suspensions are diluted to 1 million/mL in media. Suspensions are pelleted at 400g for 3 minutes and resuspended in 1 mL DPBS. The HEK suspension is treated with 1 ug/mL of Calcein Green (Thermo-Fisher) while the 3T3 suspension is treated with 2 ug/mL of Calcein Red (Thermo-Fisher) for 15 minutes at 37 degrees, followed by the addition of 4 mL media. Suspensions are pelleted and resuspended in media. Cells are mixed together in a 1:1 ratio and diluted in DPBS to form a final concentration

of 250k/mL, which contained also 10 uM Cascade Blue-Dextran (Thermo-Fisher) and 0.5 v/v% FBS are added. Jurkat cells (ATCC) are cultured in RPMI-1640 medium supplemented with 10% FBS and 1x Penicillin-Streptomycin at 37 deg and 5% CO2. 1 million cells are extracted and pelleted at 400g for 3 minutes and diluted to a in 500 uL DPBS, to which 1 uL of 5 mM DRAQ5 (Thermo-Fisher) is added. Cells are incubated at 37 degrees for 5 minutes, to which 500 uL of DPBS is added which also contained 10 uM Cascade Blue-Dextran and 0.5 v/v% FBS.

PBMC staining. Fresh-frozen human PBMCs are obtained from Stemcell Technologies. DMEM with 10% FBS is warmed up to 37 degrees, and the frozen PBMCs are thawed by adding 1 mL of warm media on top of the frozen cells and immediately transferring the media to a 15 mL conical. This process is repeated several times, and then warm media is added until the total volume in the conical was 10 mL. Cells are pelleted at 300g for 5 minutes, after which the supernatant is removed and replaced with 10 mL of warm media. Cells are counted and 500k cells are transferred into a 1.5 mL tube. After pelleting, the supernatant is removed and replaced with 1 mL ice-cold 1% BSA in PBS. Cells are pelleted, the supernatant is removed and replaced with 100 uL 1% BSA in PBS and the tube contents are transferred to a 15 mL conical. 5 uL of FcX blocking antibody (BioLegend) is added and the tube was incubated at room temperature for 5 minutes. Then, 7.5 uL of PE-conjugated CD16 antibody (BioLegend) and PE-Cy5.5-conjugated CD14 antibody (Biolegend) are added to the tube. Cells are stained at room temperature in the dark for 30 minutes. Cells are pelleted and washed twice with 5 mL 0.04% BSA in PBS, after which cells are resuspended in 1 mL 0.04% BSA in PBS, to which 2 uL of 1 mM FITC and 200 U RNAse inhibitor (Lucigen) is added. Cells are passed through a 40 um cell strainer (Fisher Scientific) prior to encapsulation in droplets.

Cell and bead encapsulation within monodisperse droplet emulsions. Barcoded mRNA capture beads are purchased through ChemGenes (MACOSKO-2011-10) and have a structure previously reported (7). Beads arrived as a dry resin and are resuspended, washed, and filtered as previously described. For each experiment, 100,000 beads are extracted from the suspension and pelleted by placing on a tabletop centrifuge for 10 seconds. The supernatant is removed and replaced with 40 uL of 10 mM 4-MU (Sigma) in methanol diluted in 960 uL DPBS. The pellet is resuspended and allowed to stain for 1 minute at room temperature. Beads are then pelleted, washed with DPBS once, then resuspended in a solution of 10 uM FITC in DPBS to which is added 500 uL of the Drop-Seq lysis buffer. Beads are then placed into a 3 mL syringe with a magnetic stir bar (V&P Scientific) and encapsulated in 2% w/v Ionic Krytox surfactant in HFE 7500 (3M) on an 80 x 80 um drop-making device. Flow rates used are 4000 uL/hr for the bead suspension and 12000 uL/hr for the oil. Cell suspension is placed in a 3 mL syringe with a magnetic stir bar and encapsulated in 2% w/v PEGylated surfactant in HFE 7500 on a 80 x 45 um drop-making device. Flow rates used are 1500 uL/hr for the cell suspension and 4000 uL/hr for the oil.

PDM operation for performing linked fluorescence and scRNA-seq analysis in nanoplates. The bead-containing droplets are passed through the PDM device at input rates of 80 - 120 Hz. Bead-containing droplets are programmably dispensed to each microwell at a maximum printing rate of 3 Hz between nanowells and 10 Hz if printing multiple beads to the same well. Following printing of beads to nanowells, cell-containing droplets are passed through the PDM device at input rates of 80 - 160 Hz. Cells are dispensed to nanowells at a printing rate of 2-5 Hz. The fluorescence of every cell printed is recorded into a text file along with its nanowell location. Following printing of cells and beads, the nanowell slide is disconnected from its power source,

causing droplets to float to the surface, where they are transferred by a P-1000 pipette into a 50 mL conical on ice.

Sequencing library preparation. The collected emulsions are processed similarly to the Drop-Seq workflow (7). In brief, the emulsion is broken, beads are collected and reverse transcribed with MMLV reverse transcriptase (Maxima RT, Thermo Fisher), unused primers are degraded with Exonuclease I (New England Biolabs), beads are washed and PCR amplified. The following modifications are incorporated to account for the low number of beads collected. During the emulsion breakage step, a 0.01% v/v solution of Sarkosyl in 6x SSC is used. During the steps leading up to reverse transcription, a 0.01% v/v Tween-20 solution in 6X SSC is used. Following PCR, the cDNA library is split into two fractions following sequential AmPure bead purification at 0.6x and 2.0x volume ratios as performed the Cite-Seq workflow. 600 pg of cDNA in the fraction containing mRNAs is processed using the Nextera XT kit to form a sequencing library. 500 pg of cDNA in the fraction containing amplified well indexes underwent a second round of PCR to add sequencing adapters. Libraries are pooled and sequenced on an Illumina machine.

NGS sequencing and optical phenotype data matching. Libraries underwent paired-end sequencing using the custom Drop-Seq primer with a read length of 25 bp for read 1 and 75 bp for read 2. For the mRNA library, reads are processed using the Drop-Seq bioinformatic pipeline. For the well index library, reads are partially processed using the Drop-Seq bioinformatic pipeline, yielding a read-quality filtered and trimmed .sam file with annotations corresponding to UMI and bead barcode positions. The CITE-seq-Count bioinformatic package is used to extract coordinate expression data from the well index sequencing library. Using this coordinate expression matrix, UMI counts for each bead are scaled based on the number of total UMIs on the bead. Next, the off-target noise of each well index is estimated based on the average

expression across all beads and subtracted from scaled UMI counts. The top x, y, and z well index captured on each bead is then extracted. Beads which the top well index is not at least 3 times as abundant as the next most abundant well index for any of the sets of x, y, and z well indexes are removed. The remaining beads are assigned to a nanowell position by matching the most abundant x, y, and z indexes on the bead to a lookup table of the expected x, y, and z positions at each nanowell position. Following position assignment, the bead barcodes of all beads matched at each nanowell position are collected. The columns on the gene expression matrix of all beads matched at the same nanowell position are merged, yielding a revised matrix where the columns represented nanowell positions instead of individual beads. The gene expression matrix is then annotated by recorded cell fluorescence values obtained during printing.

Gene expression analysis. For the cell cycle experiment, only those cells which expressed at least 300 genes and could be confidently assigned a fluorescence value are processed using the Seurat package in R. Cells are assigned a G2/M and S phase score using Seurat and a list of previous published cell-cycle associated genes (42) which is then used to assign a cell cycle state. Principal component analysis is performed using only the cell-cycle associated genes and UMAP analysis is then performed on the top 10 principal components. For the PBMC experiment, cells which expressed at least 100 genes and for which there was matching fluorescence data are selected. Gene expression is scaled and normalized using the SCTransform function within Seurat. PCA analysis is performed and UMAP clustering is performed on the first 6 principal components.

3.5 References

1. Zhao X, Guan JL. Focal adhesion kinase and its signaling pathways in cell migration and angiogenesis. 2011 Jul 18..

2. Giancotti FG. Integrin signaling: Specificity and control of cell survival and cell cycle progression. Current Opinion in Cell Biology. 1997; 9(5): 691-700.

3. Nusse R, Fuerer C, Ching W, Harnish K, Logan C, Zeng A, et al. Wnt signaling and stem cell control. In Cold Spring Harbor Symposia on Quantitative Biology; 2008. p. 59-66.

4. Hough SR, Laslett AL, Grimmond SB, Kolle G, Pera MF. A continuum of cell states spans pluripotency and lineage commitment in human embryonic stem cells. PLoS ONE. 2009 Nov 5; 4(11).

5. Shalek AK, Satija R, Shuga J, Trombetta JJ, Gennert D, Lu D, et al. Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. Nature. 2014; 510(7505): 363-369.

6. Trapnell C. Defining cell types and states with single-cell genomics. 2015.

7. Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. Cell. 2015 May 30; 161(5): 1202-1214.

8. Klein AM, Mazutis L, Akartuna I, Tallapragada N, Veres A, Li V, et al. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. Cell. 2015 May 30; 161(5): 1187-1201.

9. Fan HC, Fu GK, Fodor SPA. Combinatorial labeling of single cells for gene expression cytometry. Science. 2015 Feb 6; 347(6222).

10. Regev A, Teichmann SA, Lander ES, Amit I, Benoist C, Birney E, et al. The human cell atlas. eLife. 2017 Dec 5; 6.

11. Gry M, Rimini R, Strömberg S, Asplund A, Pontén F, Uhlén M, et al. Correlations between RNA and protein expression profiles in 23 human cell lines. BMC Genomics. 2009 Aug 7; 10.

12. Angermueller C, Clark SJ, Lee HJ, Macaulay IC, Teng MJ, Hu TX, et al. Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. Nature Methods. 2016 Feb 25; 13(3): 229-232.

13. Penter L, Dietze K, Bullinger L, Westermann J, Rahn HP, Hansmann L. FACS single cell index sorting is highly reliable and determines immune phenotypes of clonally expanded T cells. 2018 Jul 1..

14. Van Manen HJ, Kraan YM, Roos D, Otto C. Single-cell Raman and fluorescence microscopy reveal the association of lipid bodies with phagosomes in leukocytes. Proceedings of the National Academy of Sciences of the United States of America. 2005 Jul 19; 102(29): 10159-10164.

15. Lane K, Van Valen D, Defelice MM, Pe'er D, Boutet C, Covert Correspondence MW. Measuring Signaling and RNA-Seq in the Same Cell Links Gene Expression to Dynamic Patterns of NF-kB Activation In Brief. Cell Systems. 2017; 4: 458-469.e5.

16. Wilson NK, Kent DG, Theis FJ, Gö Ttgens Correspondence B. Combined Single-Cell Functional and Gene Expression Analysis Resolves Heterogeneity within Stem Cell Populations. 2015.

17. Kowalczyk MS, Tirosh I, Heckl D, Rao TN, Dixit A, Haas BJ, et al. Single-cell RNA-seq reveals changes in cell cycle and differentiation programs upon aging of hematopoietic stem cells. Genome Research. 2015 Dec 1; 25(12): 1860-1872.

18. Leng N, Chu LF, Barry C, Li Y, Choi J, Li X, et al. Oscope identifies oscillatory genes in unsynchronized single-cell RNA-seq experiments. Nature Methods. 2015 Sep 29; 12(10): 947-950.

19. Patel AP, Tirosh I, Trombetta JJ, Shalek AK, Gillespie SM, Wakimoto H, et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. Science. 2014; 344(6190): 1396-1401.

20. Prakadan SM, Shalek AK, Weitz DA. Scaling by shrinking: Empowering single-cell 'omics' with microfluidic devices. 2017 Jun 1..

21. Eng CHL, Lawson M, Zhu Q, Dries R, Koulena N, Takei Y, et al. Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH+. 2019..

22. Wang X, Allen WE, Wright MA, Sylwestrak EL, Samusik N, Vesuna S, et al. Three-dimensional intact-tissue sequencing of single-cell transcriptional states. Science. 2018 Jul 27; 361(6400).

23. Yuan J, Sheng J, Sims PA. SCOPE-Seq: a scalable technology for linking live cell imaging and single-cell RNA sequencing. Genome Biology. 2018 Dec 24; 19(1): 227.

24. Yekelchyk M, Guenther S, Preussner J, Braun T. Mono- and multi-nucleated ventricular cardiomyocytes constitute a transcriptionally homogenous cell population. Basic Research in Cardiology. 2019 Sep 1; 114(5).

25. Cima I, Kong SL, Sengupta D, Tan IB, Phyo WM, Lee D, et al. Tumor-derived circulating endothelial cell clusters in colorectal cancer. Science Translational Medicine. 2016 Jun 29; 8(345).

26. Huang H, Sikora MJ, Islam S, Chowdhury RR, Chien YH, Scriba TJ, et al. Select sequencing of clonally expanded CD8 + T cells reveals limits to clonal expansion. PNAS. 2019; 116(18): 2020.

27. Nabhan AN, Brownfield DG, Harbury PB, Krasnow MA, Desai TJ. Single-cell Wnt signaling niches maintain stemness of alveolar type 2 cells. Science. 2018 Mar 9; 359(6380): 1118-1123.

28. Cole RH, Tang SY, Siltanen CA, Shahi P, Zhang JQ, Poust S, et al. Printed droplet microfluidics for on demand dispensing of picoliter droplets and cells. Proceedings of the National Academy of Sciences of the United States of America. 2017 Aug 15; 114(33): 8728-8733.

29. Siltanen CA, Cole RH, Poust S, Chao L, Tyerman J, Kaufmann-Malaga B, et al. An Oil-Free Picodrop Bioassay Platform for Synthetic Biology. Scientific Reports. 2018 Dec 1; 8(1).

30. Fox CB, Nemeth CL, Chevalier RW, Cantlon J, Bogdanoff DB, Hsiao JC, et al. Picoliter-volume inkjet printing into planar microdevice reservoirs for low-waste, high-capacity drug loading. Bioengineering & Translational Medicine. 2017 Mar; 2(1): 9-16.

31. Kiselev VY, Andrews TS, Hemberg M. Challenges in unsupervised clustering of single-cell RNA-seq data. 2019 May 1..

32. McGinnis CS, Murrow LM, Gartner ZJ. DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. Cell Systems. 2019 Apr 24; 8(4): 329-337.e4.

33. Biočanin M, Bues J, Dainese R, Amstad E, Deplancke B. Simplified Drop-seq workflow with minimized bead loss using a bead capture and processing microfluidic chip. Lab on a Chip. 2019; 19(9): 1610-1620.

34. Brockhoff G. DNA and Proliferation Analysis by Flow Cytometry. In Brockhoff G. Cellular Diagnostics: Basic principles, methods and clinical applications of flow cytometry.: S. Karger AG; 2008. p. 390-425.

35. Becht E, McInnes L, Healy J, Dutertre CA, Kwok IWH, Ng LG, et al. Dimensionality reduction for visualizing single-cell data using UMAP. Nature Biotechnology. 2019 Jan 1; 37(1): 38-47.

36. Stansfield BK, Ingram DA. Clinical significance of monocyte heterogeneity. Clinical and Translational Medicine. 2015 Dec; 4(1).

37. Martinez FO. The transcriptome of human monocyte subsets begins to emerge. 2009..

38. Metcalf TU, Wilkinson PA, Cameron MJ, Ghneim K, Chiang C, Wertheimer AM, et al. Human Monocyte Subsets Are Transcriptionally and Functionally Altered in Aging in Response to Pattern Recognition Receptor Agonists. The Journal of Immunology. 2017 Aug 15; 199(4): 1405-1417.

39. Cross SE, Jin YS, Tondre J, Wong R, Rao JY, Gimzewski JK. AFM-based analysis of human metastatic cancer cells. Nanotechnology. 2008 Sep 24; 19(38).

40. Rubakhin SS, Churchill JD, Greenough WT, Sweedler JV. Profiling signaling peptides in single mammalian cells using mass spectrometry. Analytical Chemistry. 2006 Oct 15; 78(20): 7267-7272.

41. Macaulay IC, Ponting CP, Voet T. Single-cell multiomics: multiple measurements from single cells. Trends in Genetics. 2017 Feb 1;33(2):155-68

42. Nestorowa S, Hamey FK, Pijuan Sala B, Diamanti E, Shepherd M, Laurenti E, Wilson NK, Kent DG, Göttgens B. A single-cell resolution map of mouse hematopoietic stem and progenitor cell differentiation. Blood. 2016 Aug 25;128(8):e20-31.

3.6 Supplemental

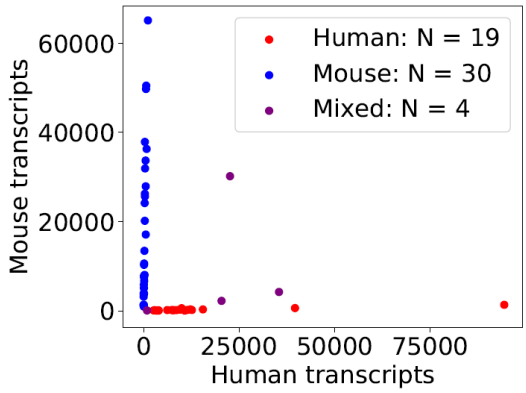

Figure S3-1. Deeper sequencing of a small subset of beads from a 2-cell experiment, showing beads that captured at least 1000 transcripts. Beads are classified as mixed beads if they captured 10% mouse or human transcripts, depending on the dominant transcript type.

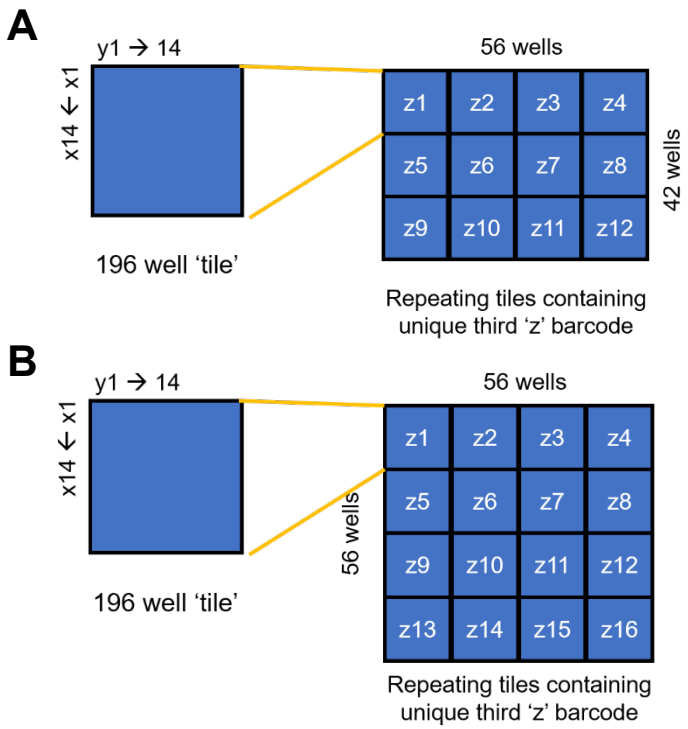

Figure S3-2. Well index coordinating scheme. Each well has a unique but known combination of three indexes. A 14 by 14 tile has each column and row barcoded with a unique index. This tile is repeated a total of (A) twelve or (B) sixteen times with each tile having a unique third index.

Table S3-1: Components of coordinate oligo, in 5' to 3' order.

Component	Sequence
TSO_PCR Primer	AAGCAGTGGTATCAACGCAGAGTGAAT
Illumina Read2 Primer	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG
Barcode	NNNNNNNN
PolyA (21) tail	AAAAAAAAAAAAAAAAAAAA

Table S3-2: Lookup table of the barcode sequences of the coordinate oligos used in this study.

Coordinate Name	8bp barcode
y1	CGCCGCAA
y2	ACCAGCCG
y3	TACTGCGA
y4	GTCGGCTG
y5	CCCAGGAC
y6	AACTGGCT
y7	TTCGGGGC
y8	GGCCGGTT
y9	CACTGTAG
y10	ATCGGTCA
y11	TGCCGTGG
y12	GCCAGTTA
y13	GGCATTGT
y14	CCCTTTTC
x1	TGCATACG
x2	GCCTTAGA
x3	CACGTATG
x4	AGCATCAC
x5	TCCTTCCT
x6	GACGTCGC
x7	CTCCTCTT
x8	ACCTTGAG

Coordinate Name	8bp barcode
x9	TACGTGCA
x10	GTCCTGGG
x11	CGCATGTA
x12	TTCCTTCC
x13	TCGTAGTG
x14	GTGCATAC
z1	GGGAAAAG
z2	CCGTAACA
z3	AAGGAAGG
z4	GCGTACAT
z5	CAGGACCC
z6	ATGCACGT
z7	TGGAACTC
z8	GAGGAGAA
z9	CGGAATCT
z10	ACGTATGC
z11	CGGTCAAT
z12	ACGGCACC
z13	TAGCCAGT
z14	GTGACATC
z15	CCGGCCAA
z16	AAGCCCCG

Chapter 4 : High diversity droplet libraries generated with a commercial liquid printer

This chapter is adapted from a manuscript in preparation with the same title by Jesse Q. Zhang, Ata Dolatmoradi, Christian A. Siltanen, Chen Sun, Zev J. Gartner, and Adam. R. Abate.

4.1 Abstract

Droplet libraries consisting of large numbers of distinct reagents encapsulated as separate droplet populations are necessary for important applications of droplet microfluidics, including combinatorial chemical synthesis, DNA-encoded libraries, and massively multiplexed PCR. However, existing approaches for generating these libraries are laborious and impractical for more than a few reagents. Here, we describe an approach to generate droplet libraries from hundreds of reagents using a commercial array spotter to sequentially emulsify reagents stored in well plates. The approach can generate a library in $1/10^{th}$ the time compared with manual operation of a microfluidic droplet generator, without any user intervention. The ease with which our approach can generate high diversity droplet libraries should make combinatorial applications more feasible. Moreover, the instrument serves as an automated droplet generator, allowing execution of droplet reactions without microfluidic instrumentation or expertise.

4.2 Introduction

The most successful implementations of droplet microfluidics share the feature that the sample to analyze consists of discrete entities, such as DNA molecules (1), functionalized beads (2,3), or cells (4,5). For example, in enzyme screening applications, the sample consists of microbes expressing variants from a library (6), while in single cell genomics, it consists of cells with unique molecular features (7). In these and similar applications, droplet analysis makes sense

because uniform reagent conditions in droplets containing different discrete entities can produce diverse results that are useful to characterize. For example, a uniform catalytic measurement can score enzyme variants by activity (8), while uniform reaction conditions can barcode and sequence thousands of single cell transcriptomes (9). There are other applications, however, in which the liquid portion of the droplets must vary. Combinatorial chemistry synthesizes chemical libraries by mixing a fixed set of input reactants in different combinations (10). Similarly, drug screens test hundreds of compounds, individually or in combination, against model systems to elicit a desired response (11). Such combinatorial applications would greatly benefit from the speed and efficiency of droplet microfluidics but are currently challenging due to the need for droplet libraries consisting of hundreds or thousands of different droplet types.

To conduct combinatorial studies the input reagents must be transformed into a droplet library that can be merged to generate combinations. A common strategy is to emulsify each solution sequentially and pool the results (11,12). Manual emulsification, however, is practical for only a few reagents. Automated emulsification is more scalable but requires custom robotics that interface with the microfluidic device, which can be technically challenging (13,14). Both approaches are limited by the speed of one drop maker, and contamination is a major issue as all solutions are emulsified by the same device. Parallel emulsification methods avoid contamination by emulsifying each solution in its own device and are much faster since hundreds of devices can operate in unison (15,16). However, parallel devices are difficult to design, build, and operate, and prone to failure due to imperfect fabrication or dust clogging some of the drop makers. The resultant emulsions often lack reagents from the library and contain large, polydispersed droplets that can make them unsuitable for further use. Because of these limitations, parallel emulsification is generally limited to libraries with few sample types. Thus, generating libraries for hundreds of

different reagents remains an unsolved problem and the major barrier to applying droplet microfluidics to combinatorial applications.

In this paper, we describe scalable generation of high diversity droplet libraries with a commercial reagent spotter. Our approach uses the Scienion sciFLEXARRAYER, an instrument designed to print droplets onto solid substrates (17). Instead, we use it to print droplets into a surfactant-laden oil bath, thereby generating stable, monodispersed droplets. Because the instrument is designed for commercial use, it has useful features for library generation, including automation, ready interfacing with well plates, and rigorous washing of dispensing surfaces between samples. The piezo-electric droplet generation is controlled, allowing size to be selected from 50-800 pL. It is fast, generating monodisperse droplets up to 500 Hz and taking a few hours to create a library of hundreds of reagents. The droplet generation is on demand, allowing the exact proportion of each reagent in the final library to be selected. The device is an automated droplet generator that can be used to perform assays typically requiring a microfluidic device and the associated expertise. As a demonstration, we use it to perform digital PCR, though other assays for enzyme screening, single cell analysis, and analyte quantitation are feasible. The speed, reliability, and ease of this approach for generating high diversity libraries should make combinatorial droplet microfluidic applications more feasible. Moreover, the ability to conduct droplet reactions without specialized microfluidic instrumentation or expertise should broaden the accessibility of these techniques.

4.3 Results and Discussion

4.3.1 A 3-step cycle for automating a commercial liquid printer for generating monodisperse droplet libraries

Commercial piezo-electric droplet printing, which normally prints onto solid substrates, can be instead used to make controlled emulsions by printing into an oil bath. Automated emulsification is accomplished using a three-step cycle repeated for each reagent (**Fig. 4-1A**). The capillary nozzle of the printer moves to the wash tray, where residual sample from previous cycles is removed; this is an essential step to minimize cross contamination. The nozzle moves to the sample plate, typically a 96 or 384 well plate containing the different solutions to be emulsified, loading the desired amount of reagent into the capillary. These standard well plates are often used to store compound libraries. Finally, the nozzle moves above the oil bath, where it ejects droplets of the reagent. The droplets pierce the oil layer and are immediately coated by surfactant, pooling beneath the oil-air interface (**Fig. 4-1B**). The library is produced at a rate of 50 - 900 µL/hr depending on droplet size and instrument settings, such that a library totaling ~5 mL can be created

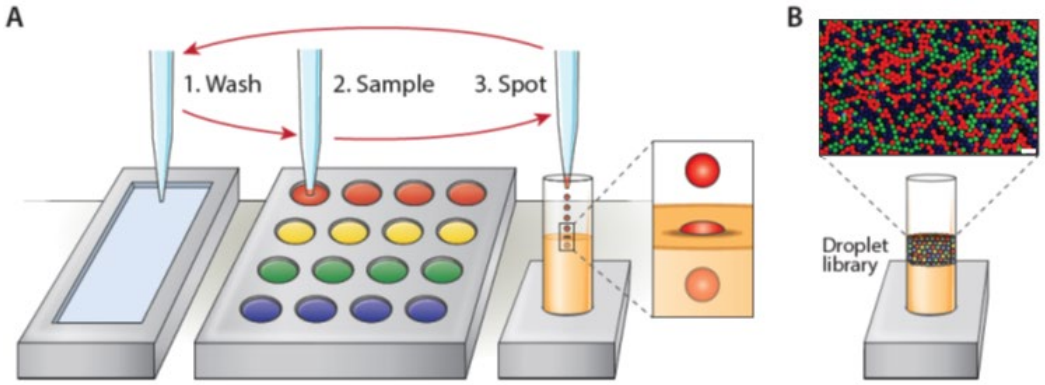

Figure 4-1. Automated liquid handling and dispensing with commercial drop in air printer for generation of monodisperse droplet libraries.
(A) Each sample is emulsified by operating the printer's capillary nozzle in a three-step cycle. 1. The nozzle moves to the wash tray and the previous contents in the nozzle are ejected; 2. The nozzle moves to the sample plate and suctions several microliters of sample; 3. The nozzle moves to the oil bath and ejects droplets into an oil bath, after which it moves to the wash tray. (B) Droplet libraries generated in this fashion are monodisperse and analogous to those generated by a microfluidic device. Scale bar = 200 µm.

in ~5.5 hr. Switching between samples adds instrument movement, capillary wash, and sampling time of about 3 min per reagent. For example, a 1 mL library of 10 samples would take ~2.5 hr to produce, while the same volume of 100 samples would take ~7 hr.

The SciFlexArrayer generates droplets on demand from static fluids maintained in the dispensing capillary via actuation of piezo-electric driven pressure pulses (**Fig. 4-2A**). Initially, the capillary is filled to the tip with the dispensing liquid; when the pulse is applied, a droplet bulges from the tip and detaches. The remaining liquid retracts up the capillary before refilling and coming to rest at the tip, where the cycle can repeat. The ejected droplet continues forward due to its inertia, piercing the top of the oil and generating ripples traveling outward from the point of impact (**Fig. 4-2B**). Once in the oil, the droplet assumes a deformed shape due to viscous drag until its inertia is fully damped and it comes to rest (**Fig. 4-2C**), at which point it floats up due to its buoyancy.

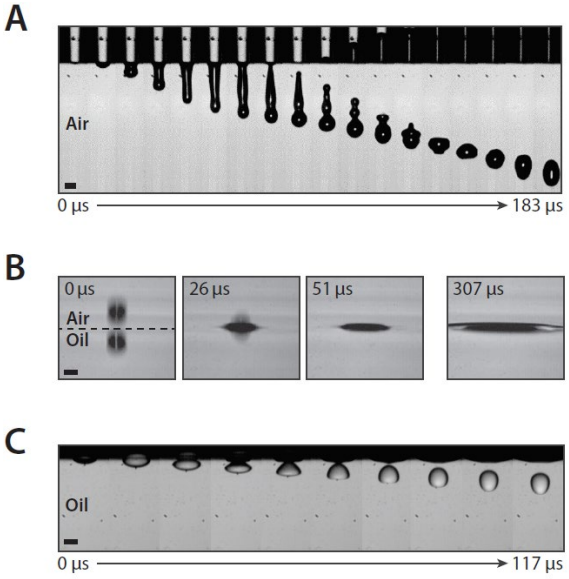

Figure 4-2. Visualization of the drop printing process.
(A) Time-lapse imagery of the droplet being ejected from the capillary tube into air by acoustic waves. (B) Impact of droplet into the oil layer. (C) Behavior of the droplet once it pierces the oil layer. All scale bars 50 µm.

4.3.2. Droplet generation is rapid, reliable, and tunable.

When operating over an exactly repeating duty cycle, this mechanism generates droplets of identical size that appear indistinguishable from microfluidic droplets (**Fig. 4-3A**). Surfactants stabilize the droplets; if omitted, droplets can coalesce in the oil. We find that 2% w/v surfactant prevents coalescence, even if droplets collide soon after entering the oil. Depending on droplet generation speed or water and oil phase composition, the surfactant may not stabilize the droplets

before collision, which could lead to coalescence. Under such circumstances, surfactant, droplet, and oil composition must be optimized to minimize coalescence. One approach to reduce coalescence is to lower the droplet generation rate to prevent sequential droplets from contacting before they are stable. For example, we find that for water droplets generated at 500 Hz monodispersity is high, but coalescence occurs at higher frequencies, resulting in polydispersity (**Fig. 4-3B, left**). We also observe that agitating the oil bath during printing by adding a stir bar reduces coalescence by reducing the chance that sequential droplets impact (**Fig. S4-1**).

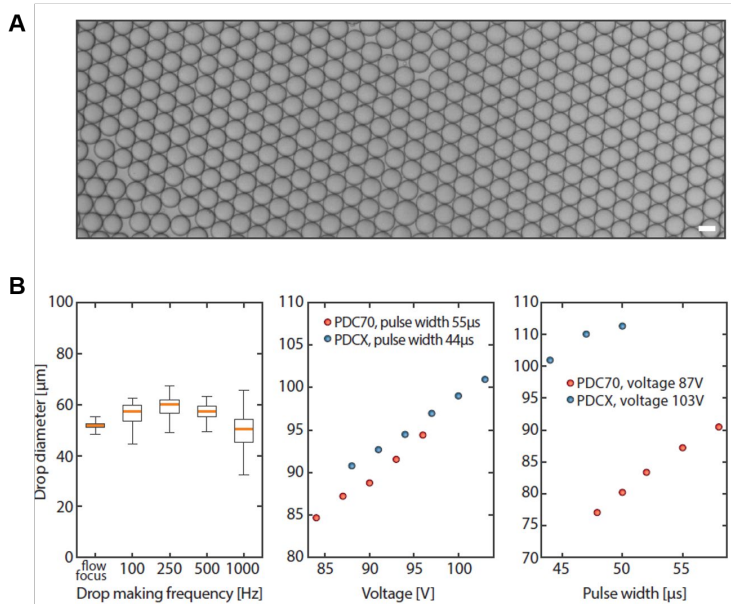

Figure 4-3. Drop-in-air printing into oil is rapid, reliable, and tunable.
(A) Micrograph of droplets generated at 300 Hz. Scale bar = 50 um. (B) Droplet size distribution as a function of acoustic wave frequency (left), voltage (center), and pulse width (right).

Since the minimum volume of an ejected droplet scales with capillary diameter, smaller capillaries generate smaller droplets. For a given capillary size, droplet diameter can be varied by tuning pulse amplitude and width. Increasing amplitude by applying higher voltages at fixed pulse width yields a linear increase in droplet diameter (**Fig. 4-3B, center**). Increasing pulse width while fixing amplitude also yields a linear increase in droplet diameter (**Fig. 4-3B, right**). Thus, tuning both parameters and using different sized capillaries affords a wide range of controlled droplet diameters. This demonstrates that the sciFLEXARRAYER is an effective instrument for generating monodispersed emulsions of controlled size without microfluidics. Thus, it should be

useful to labs interested in conducting monodispersed droplet reactions but lacking microfluidic expertise.

4.3.3. Generation and analysis of a 64-member optically encoded droplet library

In addition to generating controlled emulsions, a major value of the sciFLEXARRAYER is its ability to emulsify solutions stored in plates with full automation. To demonstrate this, we construct a reagent set comprising 3 dyes at 4 concentrations, yielding 64 combinations (**Fig. 4-4A, top**). We image the constructed emulsion in 3 fluorescence channels (**Fig. 4-4-A, bottom**). Using this image data, we extract intensity values from the center of each droplet and generate 2-D heatmaps (**Fig. 4-4B**). We generate histograms for each of the fluorescence channels (**Fig. S4-2A**) observing peaked distributions for all 4 dye concentrations. Due to crosstalk between FITC and Cy5, the distributions overlap for some droplets. To facilitate visualization of this 3D data, we perform a dimensionality reduction using T-distributed stochastic neighbor embedding (tSNE), filtering out droplets that do not cluster at the expected dye concentrations (**Fig. S4-2B**), obtaining 54 clusters (**Fig. 4-4C**). The unresolved clusters result from low intensity droplets and

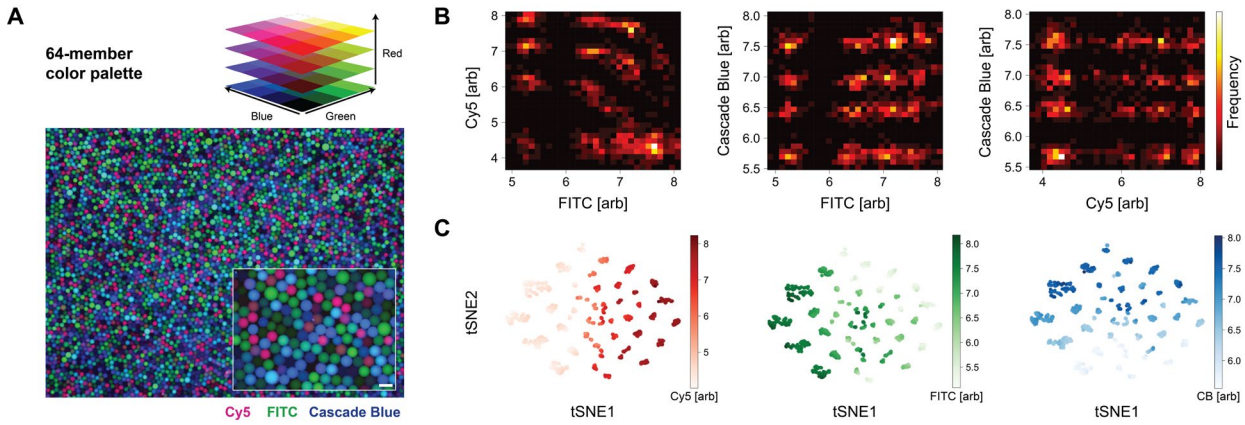

Figure 4-4. Emulsification of a large optically encoded library.
(A) A 64-member color palette of all possible combinations of 4 levels of blue, green, and red dyes is emulsified with the printer. Scale bar = 100 µm. (B) Based on images of the emulsion, brightness in the Cascade Blue (CB), FITC, and Cy5 channels is extracted from each droplet and visualized as a series of 2-D heatmaps. (C) To identify distinct droplet populations, drops are filtered and analyzed by T-distributed stochastic neighbor embedding (tSNE). The raw fluorescence data is overlaid on each analyzed droplet on the tSNE plot.

cross talk between the dyes, which become difficult to resolve by widefield fluorescence microscopy due to minor light variation at the periphery of the images (**Fig. 4-4A, bottom**). Overall, this library of ~2 mL total volume took ~4 hrs to generate from the well plate, demonstrating the effectiveness of this approach for transforming well plate libraries into droplet libraries suitable for microfluidic use.

4.3.4. Generation of a primer library in droplets

Encapsulating DNA in droplets is necessary for a broad array of applications, including digital PCR (1), enzyme screening (8), and single cell sequencing (9). To demonstrate the ability of our approach to generate oligo-containing droplet libraries that can be used for downstream assays, we generate a library comprising 192 unique primer sequences (**Fig. 4-5A**). The primer library consists of universal priming sequences flanking an 8-oligo barcode and is stored in a 384

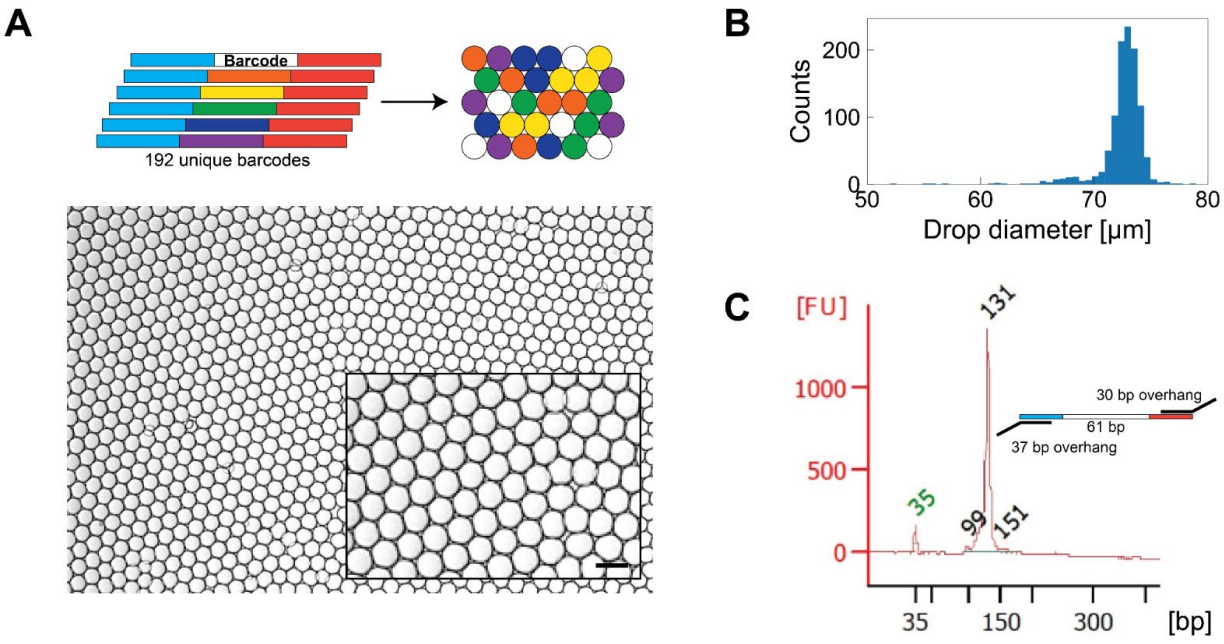

Figure 4-5. Generation of a primer library in droplets.
(A) A library consisting of oligos consisting of one of 192 barcode sequences flanked by constant regions is encapsulated within droplets. Scale bar = 100 um. (B) Histogram of the size distributions of droplets in the emulsion. (C) Encapsulated oligos are amplified by primers targeting the constant regions. DNA electropherogram confirms appropriately sized cDNA.

well plate. With conventional microfluidic techniques, generating such a 192-member droplet library would take over a day of round the clock operation with a single microfluidic device running at ~10 min per cycle, accounting for sample loading into syringes, startup of the device, and droplet generation, and would be wasteful of syringes, tubing, and labor. With our approach, we generate this library in ~8 hr without microfluidics or user intervention. The resultant emulsion is monodispersed, with a droplet standard deviation of 3.3% (**Fig. 4-5B**). To confirm the library contains functional primers, we break the emulsion and amplify the released barcodes, obtaining products of the expected size (**Fig. 4-5C**). This demonstrates that generation of a 192-member droplet library directly from a well plate is simple and effective and that the process does not harm the oligos.

4.3.5. Digital droplet PCR with printed drops

The sciFLEXARRAYER functions as an automated droplet generator that, in principle, can perform any droplet reaction compatible with well-plate storage, sampling, and the unique emulsification mechanism. Thus, it affords an accessible means by which to conduct droplet assays without microfluidic instrumentation or expertise. To demonstrate this, we use the approach to perform digital droplet PCR, a ubiquitous and important application of droplets that normally requires specialized microfluidics. As an example target we use ΦX174 virus, generating droplets at different concentrations to characterize dynamical range and accuracy. Compared to the negative control, we observe fluorescent droplets when the virus is present (**Fig. 4-6A**) obtaining a bimodal distribution corresponding to the negative and positive droplets (**Fig. 4-6B**). Because we load the virus at limiting dilution, encapsulation follows Poisson statistics (18). The Poisson estimator λ, the number of DNA molecules per droplet, thus relates to the percentage positive droplets by the following equation: $\lambda = -\ln(1 - p)$. We perform a regression on experimentally

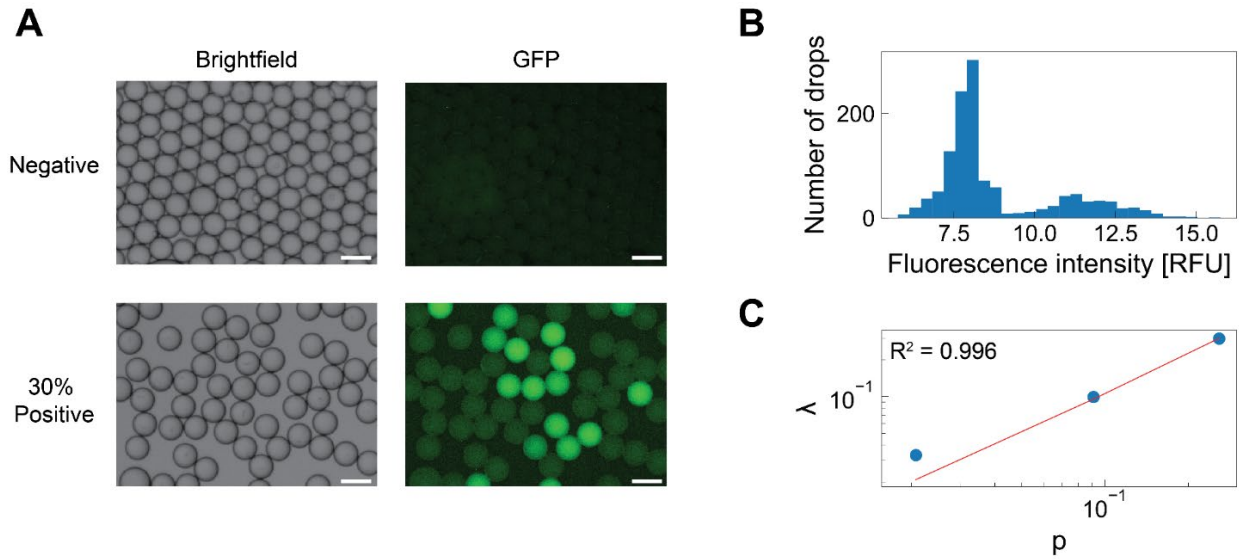

Figure 4-6. Digital droplet PCR using droplets made with the sciFLEXARRAYER.
(A) Micrographs of ΦX174 DNA co-encapsulated with digital droplet PCR reagents in brightfield and GFP channels for an emulsion without any ΦX174 DNA and where ΦX174 DNA is expected to be present in 30% of droplets. All scale bars = 100 um. (B) Histogram of fluorescence intensity within droplets for the 30% positive emulsion. (C) Poisson estimator λ as a function of the percentage of positive droplets p, with an overlaid regression.

obtained values for λ and p and find high correspondence (**Fig. 4-6C**). This shows that digital droplet PCR performed with the sciFLEXARRAYER behaves like reactions performed with conventional microfluidics. Moreover, it illustrates the promise of this approach for performing droplet reactions without microfluidic expertise.

4.3.6. Conclusion

We present a simple approach to generate large and diverse droplet libraries with full automation using a commercial liquid spotter. The instrument can encapsulate a single sample to perform droplet reactions commonly requiring microfluidics. Thus, it should be useful for labs wanting to conduct droplet reactions, but lacking microfluidic expertise. Moreover, its ability to precisely control the diameter of every formed droplet provides a unique opportunity for labeling reagents that can be used in combination with fluorescence tagging approaches. The ability of the approach to encapsulate large arrays of samples from well plates should make it useful as a

tool for manufacturing reagents required for applications of droplet microfluidics, including droplet libraries for combinatorial chemistry applications or droplet-templated particle libraries for screening and single cell analysis. Our approach makes high diversity droplet library generation simple and thus overcomes the major barrier to applying droplet microfluidics to a broad array of previously inaccessible applications.

4.4. Materials and Methods

Automated droplet library generation. A sciFLEXARRAYER S3 (Scienion AG) is used with either a PDC40, PDC70, or PDCX capillary nozzle. Prior to printing, the capillary is cleaned by exposing it to 1 mbar of oxygen plasma for 1 m in a plasma cleaner (Harrick Plasma). PBS from a source plate is aspirated into the capillary and printed into 1 mL of HFE-7500 (3M) oil with 2% (w/v) PEG-PFPE amphiphilic block copolymer surfactant (RAN Biotechnologies). Droplet size as a function of acoustic wave parameters is measured at the time of printing by automated imaging processing by a camera mounted onto the sciFLEXARRAYER. Visualization of droplet generation is performed with a Miro R311 (Vision Research) high speed camera. Collected emulsions are pipetted onto a cell counting slide and visualized on the EVOS Cell Imaging System (Thermo Fisher) and monodispersity is calculated using ImageJ. For drop size distribution comparison, a microfluidic flow-focusing device with a cross-sectional channel dimension of 40 by 40 microns is used. Syringe pumps (New Era) are used to drive 1 mL syringes (BD) filled with HFE-7500 oil with 2% (w/v) PEG-PFPE amphiphilic block copolymer or PBS into the device at flow rates of 2000 μL/hr and 500 μL/hr, respectively.

Optically encoded 64-member droplet library generation. Dextran - Cascade Blue conjugate (Thermo Fisher), Dextran - Cy5 conjugate (Thermo Fisher), and Dextran – Fluorescein conjugate (Thermo Fisher) are diluted in PBS to the following 4 concentrations: 3 μM, 15 μM, 30 μM, and

60 μM. All possible combinations of these three dyes and concentrations are mixed individually into 64 wells on a 384 well plate. A print routine is set up with a PDC70 capillary on the sciFLEXARRAYER to print 4000 drops of each mixture into 1 mL HFE-7500 oil with 5% (w/v) PEG-PFPE amphiphilic block copolymer in a 24-well plate. Collected emulsions are pipetted onto a cell counting slide and visualized on the EVOS Cell Imaging System using the DAPI, GFP, and Cy5 filter cubes (Thermo Fisher).

Fluorescence image analysis. A composite image of the three channels is cropped to include an 800-pixel diameter circle. Droplet size is analyzed and those droplets that are 2 standard deviations below the mean are excluded from downstream analyses. The intensity values at the center of each droplet in each channel is recorded. The intensity histograms for each of the 3 fluorescence channels is modeled as a mixture of 4 Gaussian distributions. Droplets with intensity values that are 1.5 standard deviation above or below the mean of the nearest Gaussian are filtered. tSNE clustering is performed with the sklearn Python package.

192-member primer droplet library. Custom oligonucleotides are ordered from IDT and kept at -20 °C until use. The 192-primer library consists of 96 sequences of the format CGGAGCTTTGCTAACGGTCGNNNNNNNNTCGTCGGCAGCGTCAGATGTGTATAAGAGACAG and 96 of the format CTTACGGATGTTGCACCAGCNNNNNNNNGTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG, where NNNNNNNN is a randomized 8 bp barcode that is a Hamming distance of at least 3 from all other barcodes. Each oligo is diluted to 5 μM in water and printed using a PDC70 into 1 mL HFE-7500 (3M) oil with 5% (w/v) PEG-PFPE amphiphilic block copolymer in a 24-well plate. A 2-mm diameter stir bar is added into the well and operated during printing. After printing, the collected emulsion is broken with an equal volume of 20% (v/v) perfluoro-1-octanol (Sigma-Aldrich) in HFE-7500. The emulsion is

amplified using 1x KAPA HiFi HotStart ReadyMix (Roche) and 1 μM of forward primer and reverse primer (IDT). The thermocycling conditions are: 95 °C for 3m; 8 cycles of 95 °C for 20s, 60 °C for 30s, 72 °C for 20s; and a final extension of 5m at 72 °C. cDNA is purified using a 2x sample volume ratio of AMPure XP (Beckman Coulter) beads and analyzed on the Agilent 2100 Bioanalyzer.

Digital droplet PCR. PhiX-174 virion DNA (New England Biolabs) is mixed with PCR reagents containing 1X Platinum Multiplex PCR Master Mix (Life Technologies), 200 nM probe (IDT), 1 μM forward primer (IDT), 1 μM reverse primer (IDT), 0.5% (v/v) Tween 20 (Sigma-Aldrich), and 2.5% (w/v) Poly(ethylene glycol) 6000 (Sigma-Aldrich). The reaction mix is printed with a PDC70 capillary into 100 μL HFE 7500 oil with 5% (w/v) PEG-PFPE amphiphilic block copolymer in a 0.2 mL PCR tube. After printing, the oil is replaced with 50 μL FC-40 oil (Sigma-Aldrich) with 5% (w/v) PEG-PFPE amphiphilic block copolymer. The emulsion is amplified using the following program on a Bio-Rad T100 thermocycler: 2m 30s at 95 °C; 35 cycles of 30s at 95 °C, 1m 30s at 60 °C, and 30s at 72 °C; and a final extension of 5m at 72 °C. The emulsion after thermocycling is imaged on the EVOS Cell Imaging System in brightfield and GFP channels. Intensity data is extracted from each droplet; coalesced droplets with a diameter greater than 80 μm were excluded from analysis.

4.5 References

1. Hindson BJ, Ness KD, Masquelier DA, Belgrader P, Heredia NJ, Makarewicz AJ, et al. High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. Analytical Chemistry. 2011 Nov 15; 83(22): 8604-8610.

2. Periyannan Rajeswari PK, Soderberg LM, Yacoub A, Leijon M, Andersson Svahn H, Joensson HN. Multiple pathogen biomarker detection using an encoded bead array in droplet PCR. Journal of Microbiological Methods. 2017 Aug 1; 139: 22-28.

3. Tang MYH, Shum HC. One-step immunoassay of C-reactive protein using droplet microfluidics. Lab on a Chip. 2016 Nov 1; 16(22): 4359-4365.

4. Zheng GXY, Terry JM, Belgrader P, Ryvkin P, Bent ZW, Wilson R, et al. Massively parallel digital transcriptional profiling of single cells. Nature Communications. 2017 Jan 16; 8(1): 1-12.

5. Pellegrino M, Sciambi A, Treusch S, Durruthy-Durruthy R, Gokhale K, Jacob J, et al. High-throughput single-cell DNA sequencing of acute myeloid leukemia tumors with droplet microfluidics. Genome Research. 2018 Sep 1; 28(9): 1345-1352.

6. Sjostrom SL, Bai Y, Huang M, Liu Z, Nielsen J, Joensson HN, et al. High-throughput screening for industrial enzyme production hosts by droplet microfluidics. Lab on a Chip. 2014 Feb 21; 14(4): 806-813.

7. Matuła K, Rivello F, Huck WTS. Single-Cell Analysis Using Droplet Microfluidics. Advanced Biosystems. 2020 Jan 26; 4(1): 1900188.

8. Romero PA, Tran TM, Abate AR. Dissecting enzyme function with microfluidic-based deep mutational scanning. .

9. Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. Cell. 2015 May 30; 161(5): 1202-1214.

10. Theberge AB, Mayot E, El Harrak A, Kleinschmidt F, Huck WTS, Griffiths AD. Microfluidic platform for combinatorial synthesis in picolitre droplets. Lab on a Chip. 2012 Apr 7; 12(7): 1320-1326.

11. Kulesa A, Kehe J, Hurtado JE, Tawde P, Blainey PC. Combinatorial drug discovery in nanoliter droplets. Proceedings of the National Academy of Sciences of the United States of America. 2018 Jun 26; 115(26): 6685-6690.

12. Brouzes E, Medkova M, Savenelli N, Marran D, Twardowski M, Hutchison JB, et al. Droplet microfluidic technology for single-cell high-throughput screening. Proceedings of the National Academy of Sciences of the United States of America. 2009 Aug 25; 106(34): 14195-14200.

13. Kaminski TS, Jakiela S, Czekalska MA, Postek W, Garstecki P. Automated generation of libraries of nL droplets. Lab on a Chip. 2012; 12(20): 3995-4002.

14. Longwell SA, Fordyce PM. micrIO: an open-source autosampler and fraction collector for automated microfluidic input–output. Lab on a Chip. 2019.

15. Ji XH, Zhang NG, Cheng W, Guo F, Liu W, Guo SS, et al. Integrated parallel microfluidic device for simultaneous preparation of multiplex optical-encoded microbeads with distinct quantum dot barcodes. Journal of Materials Chemistry. 2011 Sep 21; 21(35): 13380-13387.

16. Jeong HH, Yelleswarapu VR, Yadavali S, Issadore D, Lee D. Kilo-scale droplet generation in three-dimensional monolithic elastomer device (3D MED). Lab on a Chip. 2015 Sep 25; 15(23): 4387-4392.

17. Fox CB, Nemeth CL, Chevalier RW, Cantlon J, Bogdanoff DB, Hsiao JC, et al. Picoliter-volume inkjet printing into planar microdevice reservoirs for low-waste, high-capacity drug loading. Bioengineering & Translational Medicine. 2017 Mar 1; 2(1): 9-16.

18. Hatori MN, Kim SC, Abate AR. Particle-Templated Emulsification for Microfluidics-Free Digital Biology. Analytical Chemistry. 2018 Aug 21; 90(16): 9813-9820.

4.6 Supplemental

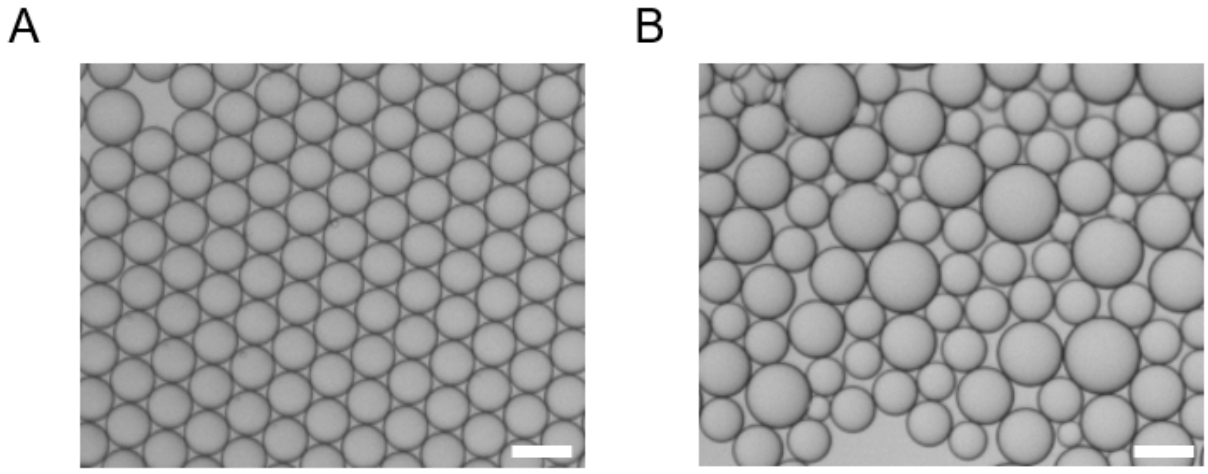

Figure S4-1. Agitating oil bath with stir bar during printing reduces coalescence. Water droplets are ejected into oil containing 2% surfactant using a PDC70 capillary with (A) or without (B) incorporation of active stirring. Scale bar = 100 um.

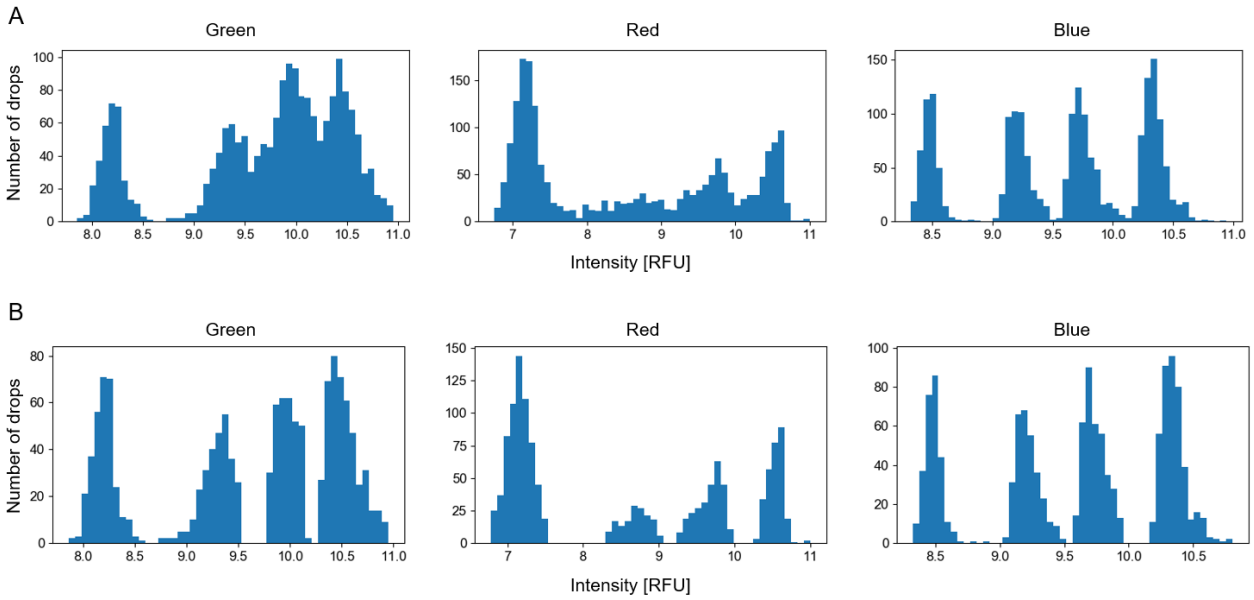

Figure S4-2. Fluorescence intensity histograms for analyzed droplets in each of the 3 imaged channels. (A) Droplet distributions without filtering. N=1958 droplets. (B) Droplet distributions after filtering out droplets in between peaks. N=1358 droplets.

Chapter 5 : A nanowell-based platform for linking single-cell RNA-sequencing with live microscopy.

5.1 Abstract

The ability to link high throughput, single-cell sequencing readouts with other phenotypic information would provide additional insights into the relationship between genotype and function at the single cell level. Microscopy and RNA-sequencing are two techniques that have established high throughput, single-cell methods and provide orthogonal, but complementary information. Existing approaches for combining single-cell RNA-sequencing and microscopy are either low throughput or high throughput but provide limited imaging information. Here, we leverage Printed Droplet Microfluidics to develop a linked imaging and RNA-sequencing assay with single-cell resolution that is compatible with standard *in vitro* culture techniques.

5.2 Introduction

In vitro cell culture systems are useful for studying phenomena observed *in vivo* in a controllable manner. By scaling down the size of the cell culture from a petri dish to a microwell, one can study a wider range of cell culture conditions at higher throughput. 96 and 384 well-plate cultures are widely employed for applications in drug discovery, gene editing, and stem cell biology (1,2,3). Micro-culture systems further scale down the size of cell culture to vessels that have length scales of hundreds or tens of microns, making single-cell or single-organoid studies feasible (4,5). Microscopy is a standard tool for assessing the phenotype of these cell cultures, including brightfield imaging to visualize morphology, fluorescence microscopy of intra- and extracellular targets, and 3D imaging with confocal microscopy. More recently, sequencing-based phenotyping has gained popularity owing to advances in high-throughput sequencing and

single-cell genomics (6). Genome-wide phenotyping for thousands of cells can be achieved with high throughput single-cell RNA-sequencing (scRNA-seq) methods (7,8), making it a popular application of single-cell genomics. Sample multiplexing techniques developed for scRNA-seq have made it possible to analyze cells from tens to hundreds of unique samples per experiment and have been applied to microculture systems (9,10). These two contrasting analysis tools provide orthogonal, but complementary information; being able to perform both sequentially on the same cells would provide additional insight not possible with either tool alone.

Several approaches exist for combined single-cell imaging and RNA-seq measurements. Manual isolation of cells using micromanipulation allows for visualization of cells before processing for RNA-sequencing in microwells (11,12). Microfluidic-based isolation tools such as the Fluidigm C1 chip enable real-time visualization of cellular dynamics immediately prior to lysis (13). Laser-capture microdissection can isolate single-cells from tissue slices, connecting *in situ* morphology to gene expression (14). However, all these approaches have throughputs of tens of cells, thereby lacking the throughput necessary for large-scale studies. Other approaches isolate and image thousands of cells in parallel within nanowells that are under 100 um in diameter (15,16). These approaches, while higher throughput, require cells to be in suspension prior to imaging. Many *in vitro* culture systems require cells to adhere to their substrate; those cells once turned into a suspension cultures lose many distinguishing aspects of their morphology. Spatial transcriptomic methods may be able to correlate *in situ* morphology with gene expression (17,18), but these methods lack true single-cell resolution. This tradeoff between resolution, information output, and throughput in existing approaches limits the capabilities of techniques that combine single-cell imaging and RNA-seq.

In this Chapter, we present a high throughput approach for linking microscopy and RNA-seq measurements at single-cell resolution that is compatible with adherent cell cultures. To accomplish this, we leverage Printed Droplet Microfluidics (PDM) and modify the nanowell plate used in PDM (19,20). By modifying the surface chemistry of the nanowell plate such that it is biocompatible for adherent cell cultures, we can perform massively parallel single-cell isolation seen with other nanowell-based imaging and scRNA-seq platforms without the limitations of utilizing cells in suspension. We demonstrate successful culture of cells within the nanowell plate and describe a workflow for extracting and barcoding transcripts from nanowells. As our approach is compatible with adherent cell cultures, richer imaging data can be obtained compared to methods imaging cells in suspension. Furthermore, our approach can be used for time-lapse imaging for phenomena that can only be observed on longer time scales.

5.3 Results and discussion

5.3.1 A workflow for culturing, imaging, and sequencing adherent cells.

The core innovation of our approach is to modify the nanowell plate used with PDM such that the individual wells are biocompatible for adherent cell culture. After fabricating the nanowell plate by photolithography, we adsorb a layer of Poly-L-Lysine (PLL) to the nanowells. This technique has been previously described to enable cells to adhere and spread compared with an untreated surface (21,22). We seed cells into the nanowell plate and incubate the cells under standard culture conditions (Figure 1A). We demonstrate that using the 3T3 cell line, cells grown within our functionalized nanowells have phenotypes that qualitatively resemble those grown on tissue culture plastic (Figure S1). Following imaging, the cells are ready to be sequenced; however, cellular mRNA quality degrades rapidly once cells are out of culture conditions. Lyophilization of mRNA is one technique previously shown in bulk cultures to preserve cellular

mRNA for weeks, even under ambient storage conditions (23). After imaging, we lyophilize the entire nanowell plate, thus preserving the mRNA in a more shelf-stable format (Figure 1B).

To connect images of cells within nanowells to sequencing data, we barcode all transcripts originating from the same nanowell. In our previously demonstrated application utilizing the nanowell plate, in which we performed linked single-cell fluorescence and RNA-seq measurements, we introduced nanowell barcodes to be able to group together reads by nanowell origin following sequencing. We introduce the nanowell barcodes in a similar fashion here, where we use a commercial liquid spotter to add coordinate oligos to bottom of each nanowell following plate lyophilization (Figure 1C). A unique coordinate oligo is spotted into each row and column, enabling sequencing reads to be traced back to specific nanowells, whereby it can be paired with a matching microscope image. To introduce mRNA capture beads into the

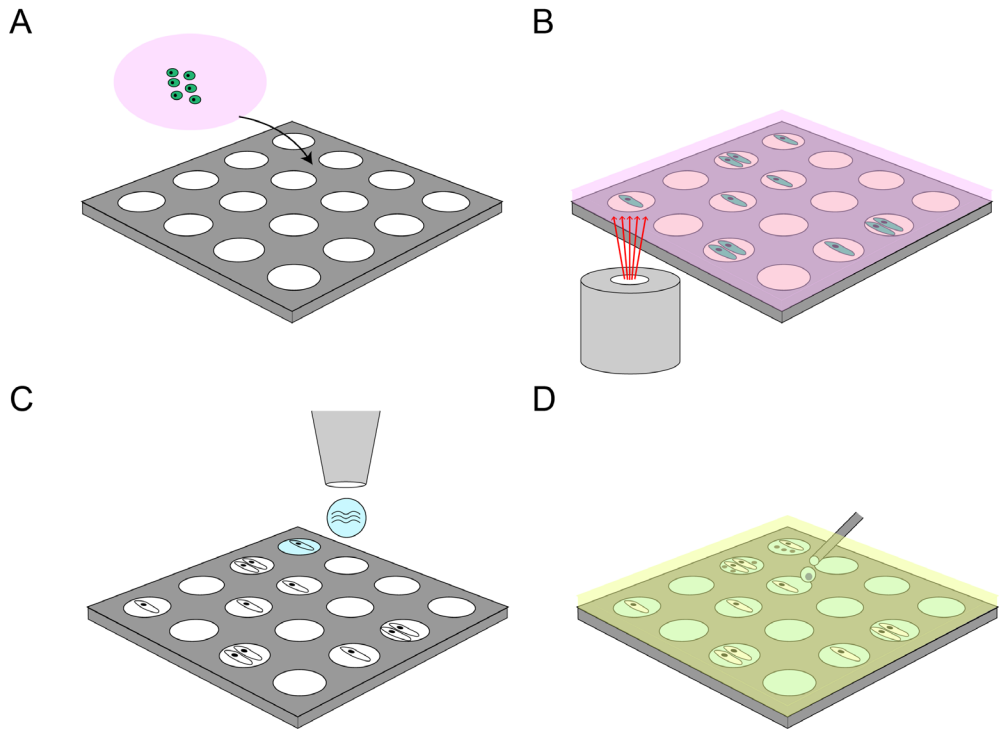

Figure 5-1. Workflow for linked microscopy and RNA-seq in nanowells.
(A) Cells in media are seeded stochastically over a nanowell plate. (B) After and during incubation, cells can be imaged with standard microscopy techniques. (C) The nanowell plate is lyophilized after imaging and nanowell barcode oligos are added with a commercial liquid spotter. (D) Cellular mRNA and barcode oligos are recovered by adding capture beads with Printed Droplet Microfluidics.

nanowells, we employ Printed Droplet Microfluidics (PDM) such that each well of interest contains beads (Figure 1D). The coordinate oligos have a poly-A tail, enabling them to be captured on poly-T mRNA capture beads along with mRNA molecules. Following printing of beads, all beads are collected and processed in parallel for sequencing.

5.3.2 Preserving cellular mRNA after imaging with lyophilization.

Lyophilization of the nanowell plate prior to PDM operation affects printing accuracy. Typically droplets are attracted and pinned to the bottom of nanowells following their ejection from the print nozzle by an attractive dielectrophoretic force. After lyophilization, a dried salt crust covers the substrate (Figure S2). These dried salts conduct away electric field lines, prevent them from extending as far above the plane of the nanowells. To wash away lyophilized salts but ensure that mRNA remains within nanowells, we employ a series of 70% Ethanol washes following lyophilization. As nucleic acids are not soluble in 70% Ethanol, this removes excess salts on the surface without also washing away the mRNA within wells. After washing the chip, printing performance is restored, allowing for robust and reliable printing of beads to nanowells (Figure S3).

5.3.3. Linked imaging and sequencing of 3T3 cells.

To determine whether our described sequencing approach is nanowell-specific, we perform a study where 3T3 cells are randomly seeded into nanowells such that each well has on average ~4 cells. Based on imaging the array after overnight cell culture, we determine the distribution of cells over the array (Figure 2A). We use PDM to add 2 beads into every well, and recover beads from 36.8% of wells, based on the beads that capture a known set of coordinate oligos. The lower capture rate is largely due to printing of fewer beads per well than in previous studies. We recover transcripts associated with those beads and map the distribution of

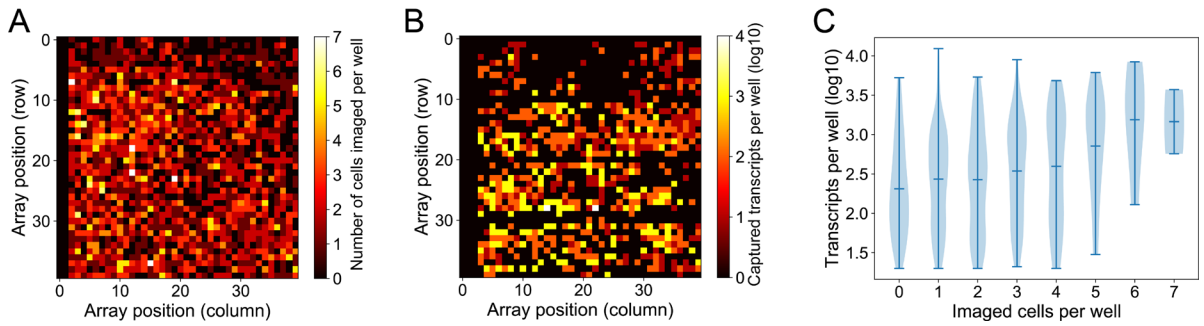

Figure 5-2. Linking of imaging and RNA-seq data for 3T3 cells randomly seeded into nanowells. (A) After overnight incubation, a 40x40 array of nanowells is imaged and the number of cells in each well counted. (B) 2 mRNA capture beads are printed to each well and recovered for sequencing. Plot of number of transcripts recovered over the 40x40 array. (C) Histogram of number of recovered transcripts as a function of the number of imaged cells in the well.

transcripts as a function of array position (Figure 2B). When matching up these two datasets, we find that for wells that recovered transcripts, 78.6% of wells contained cells by imaging. As the nanowells either contain none, a single cell, or multiple cells, we hypothesize that transcript count correlates with the number of cells within the well. When we integrate the sequencing and imaging datasets and plot transcript count as a function of the number of imaged cells for each nanowell, we find that there is a positive correlation (Figure 2C). We notice that for the false-positive wells, there is a bimodal distribution that is overlapping, suggesting that those beads that captured ambient mRNA are not always distinguishable from those beads that captured localized transcripts. Overall, we find that we can combine specific nanowell-indexed imaging and RNA-seq data.

5.3.4 Conclusion

We present a workflow for linking microscopy and RNA-seq at nanowell resolution. By lyophilizing the nanowell plate immediately after imaging, we preserve enough mRNA such that it can be sequenced using droplet-based scRNA-seq techniques. Future implementations of the method will focus on improving its fidelity such that data from a larger proportion of wells can be recovered; and that data be more reflective of the imaging data. Our method represents an

improvement over other nanowell-based implementations of imaging and RNA-seq such that it is compatible with adherent cell culture models and is amenable to time-lapse and other forms of live imaging. Our approach is modular, meaning that other single-cell resolution measurements can be integrated into this approach, such as live-imaging tracking of intracellular targets, atomic force microscopy, and mass spectrometry. The ability to connect multiple single-cell resolution measurements will further our understanding of cellular function.

5.4 Materials and Methods

Microfluidic device fabrication. Printed droplet microfluidics devices are made with poly(dimethylsiloxane) (PDMS) soft lithography as previously described (19). An 80 um by 80 um drop making device is fabricated as previously described (20).

Nanowell plate fabrication. Clean 25 mm x 75 mm glass slides have a 200 A thick layer of chromium deposited on them (LGA Thin Films). The electrode pattern on slides is patterned by positive photoresist lithography. A 2 um thick layer of MA-P 1215 (Micro Resist Technology) is spin-coated onto the slide and baked for 1 minute on a 95 degree hotplate. The slide is then exposed to collimated 190 mW UV light (Thorlabs) for 3.5 minutes. The slide is developed in MF-24A developer (Dow Chemical) for 1 minute. Photoresist is washed away with Acetone. A 15 um-thick layer of poly(dimethylsiloxane) (PDMS) is spin-cast onto the slide. The slide is then baked for 3 minutes on a 95 degree hotplate. Following plasma treatment of the slide, a 50 um thick layer of SU-8 (Microchem) is spin-cast onto the slide. The slide is soft-baked for 15 minutes on a 95 degree hotplate, then exposed to UV light under a photomask for 90 seconds, followed by 5 minutes of post-exposure baking at 95 degrees. The slide is then immersed in PGMEA developer (Sigma) for 5 minutes, rinsed with PGMEA and isopropanol, then dried on the hotplate for 2 minutes. The slide undergoes a hard bake for 30 minutes on a 200 degree hotplate. After a round of plasma treatment, the slide is immersed in 0.01% w/v poly(l-lysine) (Sigma-Aldrich) in water at 37 degrees on a shaker for 1 hour. The slide is then washed 3 times with DI water and stored in a conical with 70% Ethanol for 5 minutes. After ethanol washing, the slide is washed 2x with DI water in a biosafety cabinet, and placed in a conical with cell culture media at 4 degrees until it is ready to be seeded with cells.

Cell culture on nanowell plate. 3T3 cells (ATCC) are cultured in 75 square-cm flasks in the presence of Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and 1x Penicillin-Streptomycin at 37 deg and 5% CO2. Cells are treated with 0.25% Trypsin-EDTA and washed with media to generate cell suspensions. The viability and cell concentration are counted by a TC20 automated cell counter (BioRad). Cells are diluted to a concentration of 25k/mL. The nanowell-containing slide is incubated in media for 30 minutes prior to cell seeding. The chip is removed from media, placed in a clean petri dish. 300 uL of cell suspension is added dropwise over the slide. Cells settle for 15 minutes, after which media is flushed over the slide. The dish is placed in an incubator until imaging. Prior to imaging, cells on the nanowell plate are stained with Calcein Green (Thermo-Fisher) to aid in visualization. To achieve Calcein staining, the chip washed in PBS and immersed in a staining solution of 1 uL of 2 mM Calcein stock in 1 mL PBS and placed in a 37 degree incubator for 5 minutes. After staining, the chip is washed in PBS before replacing the PBS with cell culture media. Cells are visualized using a tile scan on a confocal microscope (Zeiss).

Lyophilization of nanowell plate. The slide containing the nanowell plate is washed 3x in PBS. A lyophilization solution containing 0.1M D-Trehalose (Sigma-Adrich) in PBS is added over the slide. The slide is dried vigorously with an air-gun then immersed into LN2 for 30 seconds to flash-freeze it. The slide is then placed in a loosely capped 50 mL conical tube and lyophilized at pressures below 50 mTorr and a temperature below -100 degrees C for at least 5 hours.

Nanowell coordinate indexing. The lyophilized nanowell plate is washed for 3 minutes 2x in 70% Ethanol that had been stored in a -20 degree freezer. During each wash the plate is placed inside of a vacuum chamber. After the second wash, the plate is dried with an air gun and placed in the vacuum chamber for 3 minutes to complete drying. Nanowell indexing then occurs as

previously described using the Scienion sciFLEXARRAYER S3 (20), except that only 10 unique x and y coordinate oligos are used to barcode a 40 x 40 array.

mRNA capture bead printing with PDM. mRNA capture beads (ChemGenes, catalog no. MACOSKO-2011-10) are stained with 4-MU as previously described. Beads are pelleted and added to 300 uL Drop-Seq lysis buffer, 600 uL PBS, and 10 uL 1 mM FITC. Beads are then placed into a 3 mL syringe with a magnetic stir bar (V&P Scientific) and encapsulated in 2% w/v Ionic Krytox surfactant in HFE 7500 (3M) on an 80 x 80 um drop-making device. Flow rates used are 4000 uL/hr for the bead suspension and 12000 uL/hr for the oil. The PDM hardware and optical configuration is identical to previously described (20). However, during printing operation, bias and spacer oils containing pure HFE-7500 (3M) are used. The nanowell plate is also immersed in a bath of pure FC-70 (3M). PDM is used to print capture beads to designated nanowells.

Sequencing library preparation. After printing, the slide is removed from the PDM printing stage and immersed in a conical containing 30 mL 6x SSC buffer with 0.01% v/v Sarkosyl (Sigma-Aldrich). The conical is shook by hand and then spun down at 1000g for 2 minutes. After spinning, the slide is removed using forceps from the conical. Sequencing library preparation follows the Drop-Seq protocol except with the modifications previously described. Libraries are sequenced on an Illumina MiSeq machine.

NGS sequencing and microscopy data matching. The sequencing library is processed as previously described (20), such that a matrix where the columns represent nanowell positions and the rows represent genes is output. The number of cells per well is estimated by the following: the brightfield image is used to crop out regions of the GFP-channel image that are in between nanowells. Then, all pixels in the GFP image are thresholded. Punctate regions

corresponding to cells are selected based on a certain threshold number of pixels. The output is a 40x40 matrix containing cell counts per nanowell. The gene expression data is converted into a 40x40 matrix where each cell contains information about the number of transcripts recovered from that nanowell.

5.5 References

1. Langhans SA. Three-dimensional in vitro cell culture models in drug discovery and drug repositioning. 2018 Jan 23..

2. Sun S, Chen C, Huang C, Zhang XD. Application of Arrayed CRISPR/cas9 Screen and its Data Analysis: A Systematic Review. In Proceedings - 2018 IEEE International Conference on Bioinformatics and Biomedicine, BIBM 2018; 2019: Institute of Electrical and Electronics Engineers Inc. p. 2549-2555.

3. Bushway PJ, Mercola M. High-Throughput Screening for Modulators of Stem Cell Differentiation. In Bushway PJ, Mercola M. Methods in Enzymology.: Academic Press Inc.; 2006. p. 300-316.

4. Vrij EJ, Espinoza S, Heilig M, Kolew A, Schneider M, Van Blitterswijk CA, et al. 3D high throughput screening and profiling of embryoid bodies in thermoformed microwell plates. Lab on a Chip. 2016 Feb 9; 16(4): 734-742.

5. Sakib S, Uchida A, Valenzuela-Leon P, Yu Y, Valli-Pulaski H, Orwig K, et al. Formation of organotypic testicular organoids in microwell culture. Biology of Reproduction. 2019 Jun 1; 100(6): 1648-1660.

6. Brazovskaja A, Treutlein B, Camp JG. High-throughput single-cell transcriptomics on organoids. 2019 Feb 1..

7. Klein AM, Mazutis L, Akartuna I, Tallapragada N, Veres A, Li V, et al. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. Cell. 2015 May 30; 161(5): 1187-1201.

8. Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. Cell. 2015 May 30; 161(5): 1202-1214.

9. Shin D, Lee W, Lee JH, Bang D. Multiplexed single-cell RNA-seq via transient barcoding for simultaneous expression profiling of various drug perturbations. Science Advances. 2019 May 15; 5(5): eaav2249.

10. McGinnis CS, Patterson DM, Winkler J, Conrad DN, Hein MY, Srivastava V, et al. MULTI-seq: sample multiplexing for single-cell RNA sequencing using lipid-tagged indices. Nature Methods. 2019 Jul 1; 16(7): 619-626.

11. Picelli S, Björklund ÅK, Faridani OR, Sagasser S, Winberg G, Sandberg R. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. Nature Methods. 2013 Nov 22; 10(11): 1096-1100.

12. Hashimshony T, Wagner F, Sher N, Yanai I. CEL-Seq: Single-Cell RNA-Seq by Multiplexed Linear Amplification. Cell Reports. 2012 Sep 27; 2(3): 666-673.

13. Lane K, Van Valen D, Defelice MM, Pe'er D, Boutet C, Covert Correspondence MW. Measuring Signaling and RNA-Seq in the Same Cell Links Gene Expression to Dynamic Patterns of NF-kB Activation In Brief. Cell Systems. 2017; 4: 458-469.

14. Foley JW, Zhu C, Jolivet P, Zhu SX, Lu P, Meaney MJ, et al. Gene expression profiling of single cells from archival tissue with laser-capture microdissection and Smart-3SEQ. 2019.

15. Yuan J, Sheng J, Sims PA. SCOPE-Seq: A scalable technology for linking live cell imaging and single-cell RNA sequencing. Genome Biology. 2018 Dec 24; 19(1): 227.

16. Goldstein LD, Chen YJJ, Dunne J, Mir A, Hubschle H, Guillory J, et al. Massively parallel nanowell-based single-cell gene expression profiling. BMC Genomics. 2017 Jul 7; 18(1): 1-10.

17. Eng CHL, Lawson M, Zhu Q, Dries R, Koulena N, Takei Y, et al. Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH+. Nature. 2019 Apr 11; 568(7751): 235-239.

18. Wang X, Allen WE, Wright MA, Sylwestrak EL, Samusik N, Vesuna S, et al. Three-dimensional intact-tissue sequencing of single-cell transcriptional states. Science. 2018 Jul 27; 361(6400).

19. Cole RH, Tang SY, Siltanen CA, Shahi P, Zhang JQ, Poust S, et al. Printed droplet microfluidics for on demand dispensing of picoliter droplets and cells. Proceedings of the National Academy of Sciences of the United States of America. 2017 Aug 15; 114(33): 8728-8733.

20. Zhang JQ, Siltanen CA, Liu L, Chang KC, Gartner ZJ, Abate AR. Linked optical and gene expression profiling of single cells at high-throughput. Genome Biology. 2020 Feb 24; 21(1): 49.

21. Wang L, Lei L, Ni XF, Shi J, Chen Y. Patterning bio-molecules for cell attachment at single cell levels in PDMS microfluidic chips. Microelectronic Engineering. 2009 Apr 1; 86(4-6): 1462-1464.

22. Jun SB, Hynd MR, Dowell-Mesfin N, Smith KL, Turner JN, Shain W, et al. Low-density neuronal networks cultured using patterned poly-l-lysine on microelectrode arrays. Journal of Neuroscience Methods. 2007 Mar 15; 160(2): 317-326.

23. Ozgyin L, Horvath A, Balint BL. Lyophilized human cells stored at room temperature preserve multiple RNA species at excellent quality for RNA sequencing. Oncotarget. 2018 Jul 31; 9(59): 31312-31329.

5.6 Supplemental

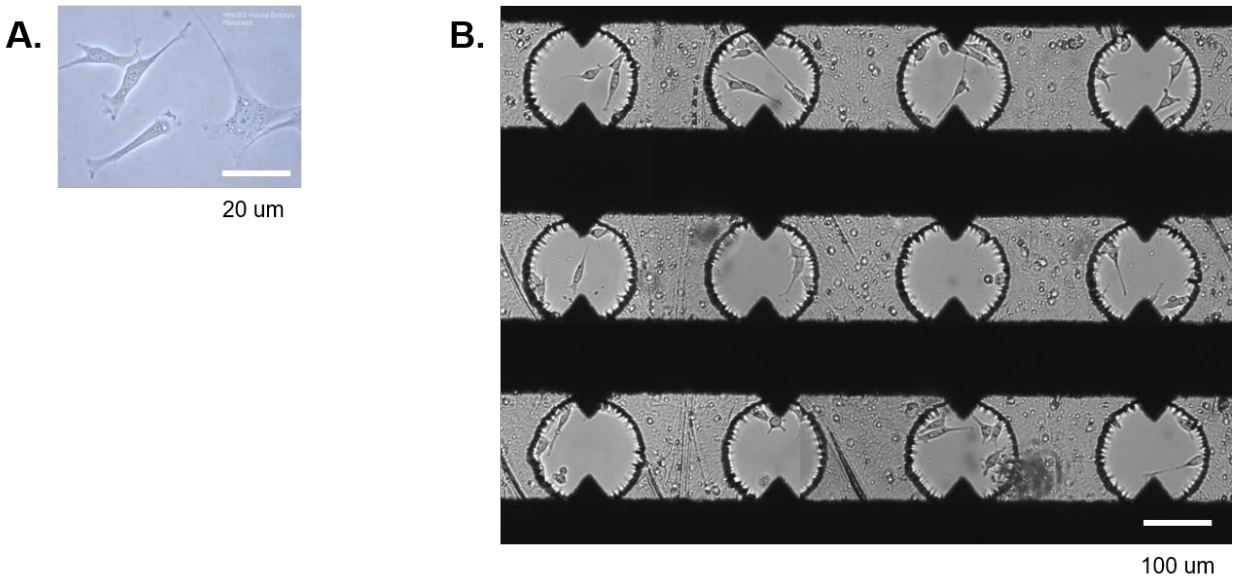

Figure S5-1. Comparison of 3T3 cell morphology. (A) 3T3 cells grown on TC plastic[1]. (B) 3T3 cells grown for 24h on a Poly-L-Lysine treated nanowell plate.

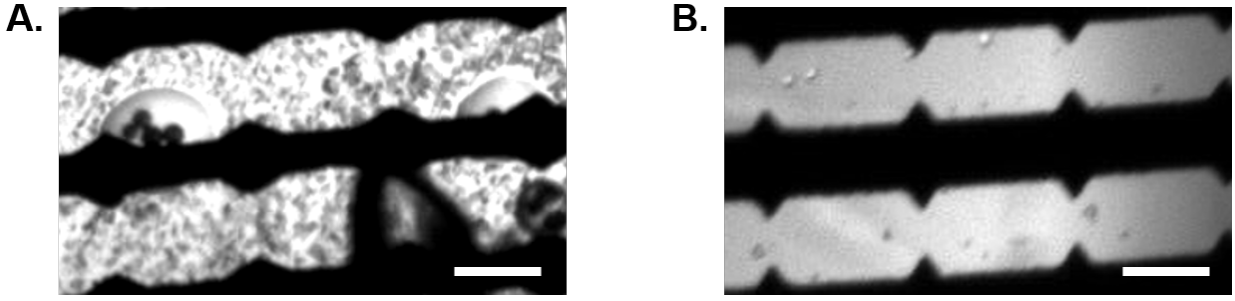

Figure S5-2. Lyophilization leaves residual salts on chip surface. (A) Image during PDM operation of printing to a chip lyophilized with PBS. (B) Image during PDM operation of a chip not subjected to[1] lyophilization. All scale bars = 300 um.

[1] Source: https://en.wikipedia.org/wiki/3T3_cells

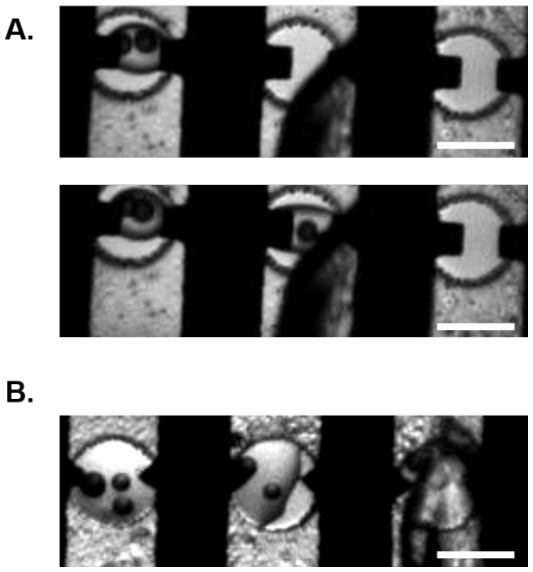

Figure S5-3. Ethanol washing recovers PDM bead printing performance. (A) Before (top) and after (bottom) stills of bead printing into nanowells on a washed chip. (B) Printing of beads into nanowells without extensive washing leads to inaccuracies and rapid wetting of printed drops. All scale bars = 150 um.

Publishing Agreement

It is the policy of the University to encourage open access and broad distribution of all theses, dissertations, and manuscripts. The Graduate Division will facilitate the distribution of UCSF theses, dissertations, and manuscripts to the UCSF Library for open access and distribution. UCSF will make such theses, dissertations, and manuscripts accessible to the public and will take reasonable steps to preserve these works in perpetuity.

I hereby grant the non-exclusive, perpetual right to The Regents of the University of California to reproduce, publicly display, distribute, preserve, and publish copies of my thesis, dissertation, or manuscript in any form or media, now existing or later derived, including access online for teaching, research, and public service purposes.

DocuSigned by:
83855EE79DE449E...
Author Signature

6/9/2020
Date