Drop-Based Microfluidics for Biological Applications

Drop-Based Microfluidics for Biological Applications

A dissertation presented

by

Yizhe Zhang

to

The Department of Chemistry and Chemical Biology

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Chemical Physics

Harvard University

Cambridge, Massachusetts

May 2015

Dissertation Advisor: Professor David A. Weitz Yizhe Zhang

Drop-Based Microfluidics for Biological Applications

Abstract

Drop-based microfluidic technology has been attracting great attention since the prevalence of soft-lithography techniques in poly-dimethylsiloxane (PDMS) microfluidic device fabrication a decade ago. The miniaturized isolated confinement of the droplet provides an ideal environment to study single cell behaviors in vitro that might otherwise be buried in the ensemble measurements. The effective confinement of the target and its secretion, together with the high-throughput processing capability, holds the promise for efficient target search through large-scale library screening. In fact, in the past seven years, considerable efforts have been made in developing this platform towards the applications in biology and great advances in drops have been reported in areas such as directed evolution, DNA sequencing, drug screening, etc.

This thesis systematically describes our work that has been done in advancing the biological application of drop-based microfluidics through three major projects that are of significance in both fundamental research and clinical applications. Encapsulating in vitro transcription and translation reactions in the 0.5 pL drops enables us to synthesize a variety of functional RNAs and proteins from the single DNA templates in a drop environment, which not only provides a novel approach for single DNA molecule detection, but also paves the way for the high-throughput screening of the artificial

proteins with drop-based microfluidics. Through successful enrichment of the restriction enzyme genes from a library consisting its truncated mutants, we demonstrated the high-throughput sorting capability of microfluidics for target gene screening that is beneficial for gene therapy applications. Finally, a less-invasive hydrogel synthesis method with microfluidic drop-maker and pico-injector is described, as a demonstration of microfluidic platform in the application of controllable synthesis of micro-sized gel particles as the 3D scaffold of, for example, Mesenchymal stem cells, for the in vitro study of cell behaviors induced by cell-cell interactions and cell-environment interactions.

Contents

List of Figures

Acknowledgements

First I would like to express my deepest gratitude to Dave, my greatest advisor. Thank him for willing to welcome me to this amazing group when I was seeking to switch my research direction from nano-materials to microfluidics six years ago. I would never forget how lucky I was in that spring to get to talk to him, learn about his group and decide on the research direction I have been working towards during the past six years. From a physicist by training in the college, to a chemistry student here at Harvard, to a biologist and an engineer through working in this group, I really got upgraded horizontally and vertically. It was never easy to span so many disciplines and try to get things to work. But it was the nonstop learning spirits I saw from Dave that kept encouraging me to take on one after another challenge, and I will keep these exploring spirits inherited from Dave in my future careers. Not only being an insightful scientist that provides us with guidance in research, Dave is also a role model to me with his unique generosity and optimistic personality. His 'sharing' attitude greatly impressed me and I believe, the entire Weitz group. I will always bear what he told us in my mind: "It is not the conservation that makes one progress. Share what you have, and if you are afraid to be left with no advantages, stay competitive by keeping improving yourself."

Due to the highly comprehensive nature of my projects, all the work to be presented here would not have been accomplished without the help and discussions from my collaborators. I therefore would like to thank the excellent biologists from New England Biolabs: Yu Zheng and Shaorong Chong for the collaborations on the restriction

enzyme project and IVT project. They have been generously providing extremely expensive reagents to me since the very beginning when I was still new to biology and probably had wasted quite a lot of samples. It was really a luxury and envied happiness to be working with scientists in bio-companies. And it was their patiently answering my one after another naïve question that helped me enter the biology field. They really can be deemed as my second advisors that intrigued my interest in molecular biology, where I am planning to keep exploring in the following years. Angelo Mao, my fellow graduate student from the Mooney lab also offered me a lot with his expertise in gel chemistry when we worked together on the stem cell niche project. His persistency in research was a good inspiration to me too. Hope him best luck in his own research.

I would also like to thank my previous and present colleagues in the Weitz lab for always being there to offer consultancy and inspirational ideas whenever I encountered problems in my experiments, for example but not limited to: Jeremy Agresti, Adam Abate, Linas Mazutis, Mark Romanowsky, Don Aubrecht, Mira Guo, John Heyman, Ralph Sperling, Pascaline Mary, Tom Kodger, Ilke Akartuna, Assaf Rotem, Ming Guo, Llyod Ung and Huidan Zhang. One thing amazing about being in such a huge research group is I was almost always able to turn to some expert for some particular question. And shaped (in a positive way) by Dave's personality, the group is always full of energy and optimism that helped cheer me up through those stressful days. I feel really lucky and proud to be a Weitz group member.

Prof. Adam Cohen from across the street is another mentor I really would love to appreciate, for his always understanding and warm support as my GAC committee all the way through my Ph.D. years. Every time meeting with him, I never failed to be inspired

from his thoughtful questions and sharp discussions. What is more invaluable, as smart as he is, Adam always keeps with him his lovely humbleness, since I first saw him eight years ago. We almost arrived here at Harvard at the same time, and I feel really happy, but not surprised, for his academic achievements! I expect to see bigger contributions from him to the scientific world!

Last but not least, the deepest thank-you to my most beloved parents back in my hometown in China. As the only child in the family, I knew how you wish me to be staying by your side just like my peers, especially during those traditional holidays. But you never mentioned a single word, just hoping me to be focusing on my work. You are my biggest motivation to strive for my goal, the strongest backup to support me to pursue my dream, and always provide the warmest comfort and encouragement in my depressing days. Most importantly, you can always see the potentials in me when no one really recognizes me. How could I let you lose!

This thesis is a summary for the past graduate research, but meanwhile is also a bright starting for the future career. Wherever I go, whatever I will be doing, I will keep the love and help I got from everyone deep in my heart for the rest of my life!

Chapter 1

Introduction

1.1 Structure of the thesis

This thesis contains three major projects, which are centered on one common theme of applying drop-based microfluidic technology to address biology related problems:

- Detecting single DNA molecules. We propose an IVT2H (in-vitro two-hybrid)-based detecting strategy on the microfluidic droplet platform to detect the unlabeled specific single DNA molecules at kHz interrogating rate. With this strategy, we successfully detected and discerned a range of model DNA molecules and therefore for the first time demonstrate the capability of drop-based microfluidics in label-free, PCR-free high-throughput detection of specific DNA at the single molecule level. (Chapter 3)
- Screening Restriction Endonucleases. We develop a novel strategy in search of Restriction Endonuclease (RE) genes based on the prokaryotic cell SOS pathway with fluorescence-activated droplet sorting (FADS) technique. With this strategy we demonstrated the successful enrichment on the target RE gene XbaI from a model library consisting of XbaI and its truncated mutant gene ΔXbaI with a series of library sizes. (Chapter 4)
- Synthesizing alginate gel beads. We develop a microfluidic pico-injection-based two-step approach to reliably fabricate control-sized alginate microgel

beads as the in vitro extracellular matrix (ECM) of the stem cells. Mesenchymal stem cells cultured in these less-invasive gel beads were viable after three days, indicating good biocompatibility of the gel fabricated with this method, and therefore a promising candidate as a 3D scaffold for long term in vitro cell culture and cell behavior studies. (Chapter 5)

1.2 Background introduction

Since the introduction of the poly-dimethylsiloxane (PDMS) soft lithography technique into the fabrication of microfluidic devices about 15 years ago [1], the studies on microfluidic devices have been surging up in that it greatly shortened the device fabrication time, reduced the fabrication complexity and therefore became amenable to more researchers as a convenient tool to perform various micro-scale experiments. In addition, the unique fluidic mechanics properties it exhibited also attracted the attention of physicists, to study the fundamental physics underlying it for a better control on the operation [2-6]. Among many other labs, our group was one of the pioneers that devoted extensive efforts to its development. For example (Figure 1.1), we demonstrated the stable production of highly monodispersed water-in-oil drops with or without cells. The cells encapsulated in the drops stayed viable and secreted antibodies off line within six hours [7]. Not just the water-in-oil drops, by modifying the channel surfaces with selective chemicals, we could also produce oil-in-water drops from several to several hundred micrometers in diameter [8-11]. By fabricating 3D devices, we could even make double emulsions and scale up the production rate [12]. Changing the geometry a bit, the drop-maker device can be applied as the drop-splitter to symmetrically or asymmetrically

break the drops for sampling purposes. Conversely, drops can also be replenished with aqueous ingredients at any time point via the operation of the pico-injector [13]. The most prominent feature of the drop-based microfluidics might be the invention and application of the sorter to detect and sort the target drops with desirable characteristics [14,15].

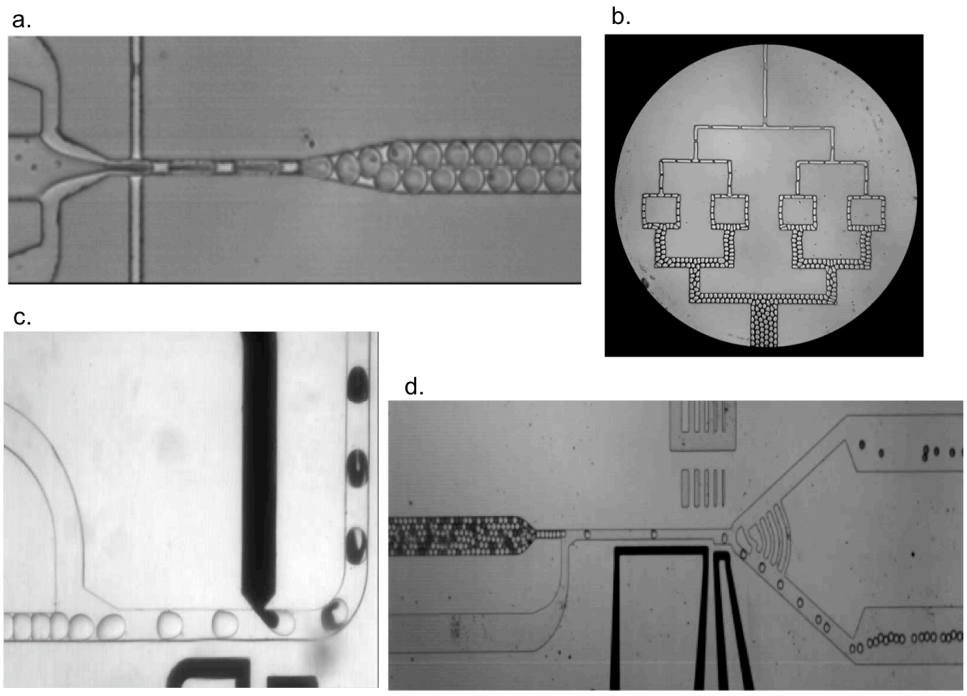


Figure 1.1 Representative micrographs of microfluidic devices on operation. (a) Co-flow encapsulation of yeast cells and the enzymatic substrate in drops on a drop-maker. (b) Symmetric drop splitter. (c) Pico-injector. (d) Sorter.

With such a highly-developed controllability and flexibility, drop-based microfluidic platform is therefore extremely suitable to be applied onto biological studies

for its pico-liter sized confinement that is ideal to include the single cell and its metabolic secretion together in an isolated microenvironment for interested cellular process detection or studies that would otherwise be buried in the noises from the environmental or neighboring cells' interruptions, and also for its ultrahigh-throughput (~ kHz) drop-interrogation capability that would enable the screening for the targets on the realistic large-scale libraries within a reasonable time frame.

In fact, there is already a report on the pioneering work of employing fluorescence-activated drop-based microfluidic sorting system to sort and direct the evolution of yeast-secreted enzyme Horseradish Peroxidase on its catalytic activity [14]. Another good example of biological applications of the drop-based microfluidics is the demonstration of sequencing DNA by monitoring the fluorescence quenching in the drops resulting from Förster resonance energy transfer (FRET) effect when a pair of DNA probes possess the perfect matching sequence to the target DNA and thus are dragged by DNA ligase to a close proximity with each other to trigger the FRET quenching. The sequence of the target DNA can hence be deducted from the known sequences of the probes [16]. These biological applications inspired me to work on my first two projects.

Chapter 2

Materials and Methods

2.1 Microfluidic Device Fabrication

The microfluidic devices were fabricated by patterning channels in PDMS using conventional soft lithography methods [1]. Briefly, for 10 μm drop-maker, 15 μm drop-maker, 25 μm pico-injector and 25 μm sorter that were used in our experiments, SU8-3010, SU8-3015 and SU8-3025 photoresists (MicroChem Corp.) were spin-coated onto the 3" silicon wafers respectively and patterned by UV exposure through a photolithography mask. After baking and development with SU-8 developer (propylene glycol methyl ether acetate; MicroChem Corp.), the 10 μm, 15 μm and 25 μm tall positive masters of the devices were formed on the silicon wafers. Then a 10:1 (w/w) mixture of Sylgard 184 silicone elastomer and curing agent (Dow Corning Corp), degassed under vacuum, was poured onto the masters and cured at 65 °C for 2 hours. Afterwards, the structured PDMS replicas were peeled from the masters and inlet/outlet ports were punched out of the PDMS with a 0.75 mm-diameter biopsy punch (Harris Unicore). The PDMS replicas were then washed with isopropanol, dried with pressurized air, and bonded to 50 × 75 mm glass slides (VWR) using oxygen plasma treatment to form the devices.

For the pico-injector and sorter, to fabricate the electrodes into the device, a 0.1 M solution of 3-mercaptotrimethoxysilane (Gelest) in acetonitrile (99.8%; Sigma) was

flushed through the electrode channels and blown out with pressurized air. A low melting point solder (Indalloy 19 (52 In, 32.5 Bi, 16.5 Sn) 0.020" diameter wire; Indium Corp.) was introduced into the electrode channels at 80 °C, followed by an eight-pin terminal block with male pins (DigiKey) glued with Loctite 352 (Henkel) to the surface of the device for strain relief. The solid electrodes in the shape of the channels were then formed when the device was cooled to the room temperature. Electrical contacts were made with alligator clips connected to a high voltage amplifier (Trek) and the function generator from the FPGA (field-programmable gate array) card running on the custom LabView program (National Instruments).

To enable the formation of aqueous-in-oil emulsions, the microfluidic channels were treated hydrophobic by flushing Aquapel (PPG Industries) into the device channels and immediately drying with pressurized air. For the drop-maker to be used in in vitro protein synthesis, a special PDMS-functionalization needs to be applied to avoid the wetting due to the existence of considerable biomolecule components in our aqueous phase. Instead of Aquapel, we passivated the drop-maker with a fluorosilane solution (1% (v/v) of 1H,1H,2H,2H-perfluorooctyl-trichlorosilane (Sigma) in Novec HFE 7500 oil (3M)) by flushing it into the device and immediately drying the channels with pressurized air.

To stabilize the drops against coalescence, we used EA surfactant donated by RainDance Technologies. The surfactant was dissolved in the fluorinated carrier oil HFE 7500 at a concentration of 1.8% (w/w).

2.2 In Vitro Protein Synthesis (IVPS) Reaction Preparation

All the DNA constructs used in this work were designed and synthesized as reported [17]. The PURExpress® in vitro synthesis kit (New England Biolabs) was used for in vitro protein synthesis. The manufacturer-recommended DNA concentration in single-expression bulk IVPS is 25-1000 ng/25 µl. To ensure the maximal production of the reporter protein in the multi-protein expressing IVPS, we controlled the amount of the target DNA slightly less than the reporter DNA. Specifically, reporter DNA was kept at 6 ng/µl while target DNA was diluted to be 3 ng/µl, 0.3 ng/µl, 0.24 ng/µl, 0.15 ng/µl and 0.03 ng/µl for the experiments at λ (average DNA molecule number per droplet) = 10, 1, 0.8, 0.5 and 0.1. We prepared the aqueous phase with 20 µl solution A (including necessary enzymes), 24 µl solution B' (including small molecules and 300 ng reporter DNA), 2 µl target DNA and 4 µl Nuclease-free H_2O, mixing them well for encapsulation into 10 µm-diameter droplets. For the detection of DNA *cro-ncoa1*, where partner DNA *er-ad* was introduced in the IVPS, 2 µl *er-ad* (3 ng/µl) was added and balanced by water to reach a total volume of 50 µl. All the reagents were handled on ice before emulsification to prevent the IVPS initiation in bulk.

2.3 Restriction Endonuclease Screening Preparation

2.3.1 Host cell preparation

ER 2745, an Escherichia coli (E. coli) BL 21 strain that contains a fusion of a DNA damage-inducing promoter and an indicator gene dinD1::lacZ, was provided by our collaborator from New England Biolabs (NEB). The strain is deficient in all known endogenous restriction systems and expresses T7 RNA polymerase under lac control from a chromosomal location. ER 2745 was transformed with plasmid pLysY on the

pACYC/Chloramphenicol (Cmp) vector that encodes for the inhibitor LysY to reduce the base level of T7 RNA polymerase expression. ER 2745/pLysY was then transformed with pTXB1_XbaI ("+" cell) or pTXB1_ΔXbaI ("-" cell) for the model library construction.

2.3.2 Model gene library design

A model library consisting of the mixture of "+" cells and "-" cells was constructed to prove the principle of the screening strategy. The quantities of the "+" cells and the "-" cells were adjusted to the ratios of 1:9, 1:99 and 1:999 respectively for the single round sorting demonstrations.

The primer p_{fwd} (5'-TAGGGGAATTGTGAGCGGATAAC-3') and p_{rev} (5'-GGAATCGGCCCTTGTTTTGATAG-3') was designed to amplify a fraction of the target gene (XbaI, 264 bp) and the control gene (ΔXbaI, 208 bp) for enrichment analysis on the agarose gel electrophoresis results on the sorted samples from droplet experiments.

2.3.3 Cell culture

Host E. coli cells from an overnight culture in LB/Amp+Cmp (Sigma) medium (Lysogeny Broth medium containing the antibiotics Ampicillin and Chloramphenicol) were allowed to grow till the OD_{600} reached around 0.2 for Isopropyl β-D-1-thiogalactopyranoside (IPTG; Life Technologies) induction.

2.3.4 Bulk fluorescence measurement

IPTG was added to the cell culture for a final concentration of 0.5 mM to induce the target gene expression. After 30 min induction, 1 ml cell culture was taken for sonication after washed into the 0.7 ml sonication buffer (100 mM NaCl (Sigma), 25 mM TrisHCl (Sigma) and 10 mM β-mercaptoethanol (β-ME; Sigma); pH 8.0) to extract β-galactosidase (β-gal). 5 μl of the cell lysate was added to the individual well of a 96-well plate that was filled with 50 μl of the substrate solution (0.2 mM Fluorescein-Di-β-D-Galactopyranoside (FDG; Life Technologies)) for LacZ activity assay. Fluorescence (EM 490/AB 514 nm) was measured after a 20-min incubation at 37 ˚C.

2.3.5 Microfluidic operations

We separated the cell suspension and the inducer before drop-making process using the co-flow strategy to prevent the induction of the β-gal expression in bulk. Briefly, cells were suspended at the density of 10^8 cells/ml (for the average of 0.3 cells per 23 μm diameter drop) in the LB solution (34 μg/ml Cmp, 100 μg/ml Amp, 0.001% Pluronic F127 (Sigma)) for inner channel infusion. Mixture of IPTG (1 mM), FDG (200 μM), N-Lauroylsarcosine sodium salt solution (0.1 %; Sigma), Pluronic F127 (0.001 %), Cmp (34 μg/ml) and Amp (100 μg/ml) in LB was introduced into the middle channel of the 10 μm drop-maker. Fluorinated oil HFE 7500 with 1.8 % EA surfactant was infused into the outer channel. To produce the 23 μm diameter drops, the flow rates for the inner, middle and outer phase are 19, 19 and 20 μl/h respectively.

The emulsion was collected in a 1 ml plastic syringe (BD) and incubated at 37 ˚C in dark for 3h to allow sufficient enzymatic reactions before re-injected into a 25-μm sorter, where the evenly spaced drops (20 μl/h) separated by carrier oil HFE 7500 (200

μl/h) were interrogated in single file at the detection window for their fluorescence at the frequency of ~1.5 kHz. The detected fluorescence signal above the certain threshold would trigger on the single-ended electric square waves (0.8~1.2 kV_{pp}, 5 cycles of 20 kHz) applied onto the electrodes that were aligned at the sorting junction of the sorter and therefore deflected the target drops into the collecting channel via the dielectrophoretic force. The asymmetric design on the downstream of the sorting junction determined the default trajectory of the drops into the waste channel when the electric field was not activated [14,15]. Fluorescence signal was acquired and processed by photomultiplier tube (PMT; Hammamatsu). The feedback loop of the fluorescence detection and electric wave generation was accomplished through the FPGA card running on a custom LabView program.

2.3.6 Colony PCR procedure

Collected samples from the microfluidic sorting were washed in Perfluorooctane solution (20% (v/v) in HFE 7500; Sigma) to break the emulsion. Cells were retrieved in 50 μl nuclease-free water (Life Technologies) for colony PCR, where 1.25 μl of p_{fwd} (10 μM) and p_{fwd} (10 μM), 0.5 μl of dNTPs (10 mM; Life Technologies), 0.25 μl of Phusion Hot Start DNA polymerase (2 unit/μl; NEB), 5 μl of 5x Phusion buffer (NEB), 2.5 μl of the cell suspension and 14.25 μl of nuclease-free water were mixed up gently for a 25 μl reaction to be undergoing a 60-cycle PCR process (initial denaturing: 98 ˚C 3 min; denaturing: 98 ˚C 10 sec; annealing: 59.2 ˚C 30 sec; extension: 72 ˚C 1 min).

2.3.7 Gel electrophoresis

1 µl of DNA amplicons from colony PCR was mixed with 1 µl Gel Loading Dye (6x; NEB) and 4 µl nuclease-free water for each well on the 1.2 % agarose (Sigma) gel with Ethidium Bromide (1 mg/µl; Life Technologies). 1 µl of 2-Log DNA Ladder (200 µg/ml; NEB) was mixed with 1 µl Gel Loading Dye (6x) and 4 µl nuclease-free water for the ladder well. The gel electrophoresis was running in 0.5x TBE buffer (Life Technologies) containing 0.5 mg/µl Ethidium Bromide at 70 V for 45 min.

2.4 Mesenchymal Stem Cell Gel-Encapsulation Preparation

2.4.1 Alginate solution preparation

Alginate (Protanol LF 20/40; FMC Technologies) was functionalized with RGD peptide GGGGRGDSP of which 10% was Fluorescein isothiocyanate (FITC)-labeled for a uniform fluorescence under the confocal microscope (Peptides International) using standard carbodiimide chemistry as reported [18]. Briefly, for the DS (Degree of Substitution) 5 modification, 1 g of sodium alginate was dissolved to 1% (w/v) in a 0.1 M MES (Sigma) and 0.3 M sodium chloride (Sigma) buffer solution at pH 6.5, followed by the addition of 68.5 mg of sulfo-NHS (Pierce), 121 mg of N-(3-Dimethylaminopropyl)-N'-ethylcarbodiimide (EDC) (Sigma) and 28 mg of RGD peptide. The reaction was allowed to proceed for 20 h before being quenched by the addition of hydroxylamine (Sigma). The reaction mixtures were then dialyzed against a decreasing concentration of sodium chloride for 2-3 days to remove salts and any unbound peptide, followed by the de-coloring step with activated charcoal. Finally the alginate was sterile (0.22 µm) filtered, lyophilized and reconstituted in serum-free DMEM (Dulbecco's Modified Eagles Medium; Life Technologies) for ionic crosslinking.

2.4.2 Stem cell suspension preparation

Clonally derived murine mesenchymal stem cells (mMSCs) (D1s; ATCC) were cultured in DMEM supplemented with 13 mM HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; Life Technologies), 2 mM Sodium Pyruvate (Life Technologies), 10% FBS (Fetal Bovine Serum; Life Technologies) and 1% penicillin/streptomycin (Life Technologies) in a 37°C, 5% CO_2 environment. The culture medium was refreshed every three days. To prepare the single-cell suspension suitable for infusion into microfluidic devices, the cells were trypsinized (Trypsin; Life Technologies), centrifuged at 1400 rpm for 5 min, and re-suspended into Dulbecco's Phosphate Buffered Saline (dPBS; Life Technologies). The PBS wash was repeated a second time to remove unbound proteins. Cells were then re-suspended in serum-free DMEM/HEPES/Sodium Pyruvate for the final concentration of about 2×10^7 per ml.

2.4.3 Microfluidic operations

All the microfluidic devices, connecting tubing (Fisher Scientific), plastic syringes (BD), needles (BD), collecting tubes (Eppendorf), pipette tips (VWR) were sterilized with UV illumination. A flow of 70% ethanol through the tubing, syringes and needles was applied before each round of experiment to prevent the bacterial contamination. PhD 2000 syringe pumps (Harvard Apparatus, Inc.) were used to infuse the fluids into the device.

To produce the cell-encapsulated alginate drops, freshly prepared D1 mMSC suspension was infused into the inner channel of the 15 μm drop-maker at a flow rate of

60 µl/h. The 2.5% (w/v) alginate solution was co-flowed in the middle channel at the flow rate of 60 µl/h to form the cell-alginate mixture at the first junction of the device before sheered into the droplet at the second junction by the fluorinated oil HFE 7500 that contained EA surfactant in the outer channel with the flow rate of 240 µl/h.

To produce the cell-encapsulated alginate drops, freshly prepared D1 mMSC suspension was infused into the inner channel of the 15 µm drop-maker at a flow rate of 60 µl/h. The 2.5% (w/v) alginate solution was co-flowed in the middle channel at the flow rate of 60 µl/h to form the cell-alginate mixture at the first junction of the device before sheered into the droplet at the second junction by the fluorinated oil HFE 7500 that contained EA surfactant in the outer channel with the flow rate of 240 µl/h.

2.4.4 Gel-cell scaffold exaction from the oil phase

After the gelation process, the collected emulsion was washed with 20% (v/v) Perfluorooctane (PFO; Sigma) in HFE 7500 solution to remove the oil phase and the cell-gel pellet was re-suspended in the serum-free DMEM supplemented with HEPES and Sodium Pyruvate for long-term culture.

2.4.5 Cytotoxicity test on the cells cultured on the synthetic extracellular matrix (ECM)

Trypan blue based exclusion assay was employed to assess the cell viability cultured on the in vitro ECM. A 1:1 (v/v) mixture of the cell-gel suspension and the sterile (0.22 µm) filtered trypan blue solution (0.4% trypan blue in 0.81% sodium chloride and 0.06% dibasic potassium phosphate; Life Technologies) was prepared and

incubated for 2 min at room temperature for viable/non-viable cell imaging under the phase contrast microscope. In some imaging, cell membrane dye PKH26 (Sigma) was employed for a better identification of the cells with regard to the scaffold gel beads.

Chapter 3

High-Throughput, Label-Free Single Molecule Detection of Specific DNA with Drop-Based Microfluidics

3.1 Introduction

Single-molecule level studies on biological processes are crucial in revealing the information that might otherwise be buried in the ensemble measurements [19]. A large number of breakthroughs in fundamental research such as DNA folding kinetics [20-22], protein-DNA interactions [23-25], DNA replication [26-32] and molecular transport [33,34], have benefited from the development in single molecule techniques. However, most of the sensitive methodologies used for single-molecule detection of DNA so far, for example, fluorescence/Förster resonance energy transfer (FRET), fluorescence correlation spectroscopy (FCS), total internal reflection fluorescence microscopy (TIRF) and surface-enhanced Raman spectroscopy (SERS), require the labeling of the target DNA [35,36], which might interfere with the native properties to be detected. For example, labeling DNA with fluorescent dyes, such as YOYO-1, was recently found to elongate and twist the intrinsic structure of DNA and change its charge property [37,38]. Therefore, lots of efforts have been made in developing label-free techniques to realize single molecule detection of DNA.

Atomic force microscopy (AFM) is a powerful label-free technique for manipulating and detecting single DNA molecules, however, it is limited by the scanning speed, and the scanning probe may cause the perturbation to the target DNA [39-41]. Another extensively investigated label-free detection scheme is based on the conductance change induced by DNA in the electric circuit composed of semiconductor nanomaterials when the target DNA is hybridizing to the complementary probe covalently-linked to the surface of the detector or translocating through a nanopore fabricated in the device [42-44]. These approaches could achieve a high detecting sensitivity as well as good temporal resolution that are critical in studying the kinetics of some microsecond-scaled biomolecular processes [44]. But the costly and complicated device-fabrication steps put a limit on its application. Although some recent reports introduced a convenient approach in fabricating glass nanopores embedded in the microfluidic device for single DNA molecules detection [45,46], it is unable to distinguish and detect the sequence-specific target; and like all the nanopore-based detection schemes, frequent clogging of the nanopore poses a big problem in practical applications.

In this report, we propose an IVT2H (in-vitro two-hybrid)-based detecting strategy on the drop-based microfluidic platform to detect the unlabeled specific single DNA molecules at kHz interrogating rate. Specifically, the target DNA serves as the template for the in-vitro protein synthesis (IVPS) of the transcription factor-like protein (target protein) that will activate the expression of the reporter-a fluorescent protein in the droplet. Therefore, by detecting the signals from the reporter, we are able to detect the presence of the specific DNA as sensitively as at the single molecule level. Actually, the picoliter drops can effectively suppress the volume-proportional background noises from

impurities and solvent in DNA solution [47,48], while in-vitro transcription and translation in drop help transmit and amplify the otherwise weak signals from the single DNA molecules without the need of labeling. With this strategy, we successfully detected and discerned a range of model DNA molecules utilizing the intermediate protein's interaction with the binding site on the reporter gene construct. We further challenged the system by applying it to the detection of a target DNA, whose protein (target protein) as a transcription regulator has to interact with another intermediate protein (partner protein) to activate the transcription and translation of the reporter gene. The success in the detection at the single molecule level further demonstrates the sensitivity and robustness of our detecting system and enables a broader range of DNA to be detectable with this assay.

3.2 Results and Discussion

3.2.1 DNA detection strategy on the drop-based microfluidics

The premise of our drop-based DNA detection strategy is the fact that one DNA codes for an increased number of its protein through repeated transcription and translation processes. This provides an initial signal transduction and amplification. Benefited from the isolated confinement provided by drops, the expressed protein can continue to trigger, as an activator, the expression of a second gene in the droplet and therefore lead to a secondary signal transduction and amplification from the target DNA. And this avalanche signal transduction and amplification process can keep on proceeding until a detectable signal level is achieved. It is thus crucial to design a tight gene expression regulation mechanism that could reliably connect the target DNA with the

reporter protein, via the intermediate protein products and their molecular interactions, for which we choose adapted transcription regulation machinery from Escherichia coli σ^{54} promoter activating system realized in the IVPS environment [17].

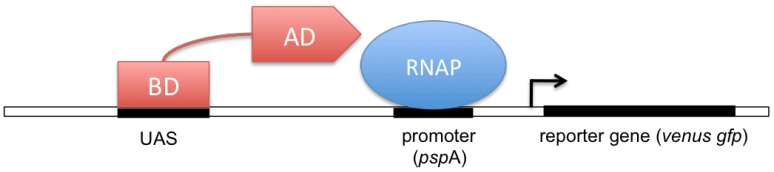
The gene expression regulation construct designed for signal transduction and amplification is based on the two-hybrid assay proposed in yeast and bacteria models [49-52], where the weak promoter-controlled RNA polymerase (RNAP) is unable to effectively initiate the transcription of the reporter gene unless the hybrid activator pair has a strong interaction and binds to the upstream activating sequence (UAS) adjacent to the promoter. In our adapted scheme, the enhancer binding protein PspF that activates the transcription of the gene under σ^{54} promoter *psp*A in E. coli is the prototype of our synthesized activator. It binds the enhancer UAS with the DNA binding domain (BD) and activates the transcription with its activating domain (AD) through interacting with the inactive *psp*A-bound σ^{54} -RNAP complex (Figure 3.1a). It has been verified in previous work that replacing BD of PspF and its enhancer with other DNA binding proteins and the recognition site does not affect the activation qualitatively [17], since BD and AD function independently. So we fuse our target DNA with the DNA encoding AD of PspF (*ad*) to form the hybrid protein as the synthetic activator to realize the DNA detection. For the proteins that don't specifically bind to DNA, their DNAs can be detected through a further modified two-hybrid strategy (Figure 3.1b), where the target DNA x is fused to the DNA for BD (*bd*) and its partner DNA y, whose protein has a specific binding to the target protein is fused to *ad*. Furthermore, an intermediate RNA with binding sequences to both X and Y can be introduced for the detection of the DNAs whose protein interacts with RNA, instead of DNA or proteins. Since most of DNAs of interest in biological

research have specific binding to other biomolecules for their physiological functions, this molecular interaction-based DNA detection scheme therefore has vast and practical applications. Additionally, the reported correlation between the transcription activation level and the molecular interaction strength in both cell-based and in-vitro two-hybrid experiments [17,53] enables our drop-based detection approach to quantitatively measure the molecular interactions in a high-throughput means alongside the DNA sensing, which has not been reported in conventional single molecule detection techniques.

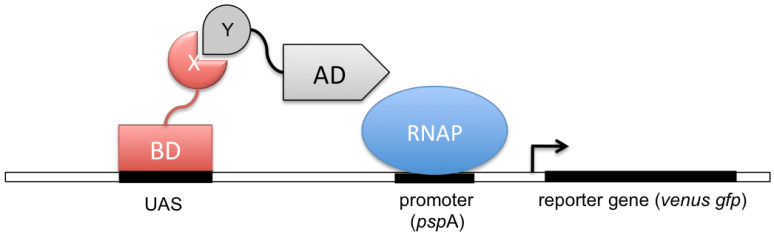
We used venus GFP as the reporter protein for its relatively simple structure compared to commonly used fluorescence-inducing proteins such as β-galactosidase (β gal). The verification of the in-vitro-hybrid-based adapted E. coli σ^{54} promoter activating system has been conducted in bulk [17]. And the cell-free single-type-protein synthesis experiments in drops have also been demonstrated on proteins EGFP, laccase, and β-gal [54-56]. Since in our system, multiple functional-related proteins are to be expressed cascade in one IVPS drop, more details have to be taken into consideration in the experimental design to optimize the synthesis conditions for a maximal production of the reporter protein, which is extremely crucial in our single molecule detection experiments. Among many adjustable factors, the ratio of the DNA amount between the target gene and the reporter gene is especially critical, in that those two genes are under different promoters (a weak E. coli σ^{54} promoter for the reporter gene, and the strong T7 promoter for the target gene). Because of this difference in transcription efficiency, when synthesized altogether under a resource-limited condition as in IVPS, there is a competition between the two genes for synthesizing resources such as enzymes, ATPs and small molecules, therefore leading to a possible synthesis bias to the target protein.

To avoid that, we finely tuned the ratio of the DNA amount between the target gene and the reporter gene and performed the IVTS in drop within the optimized ratio range.

a.



b.



c.

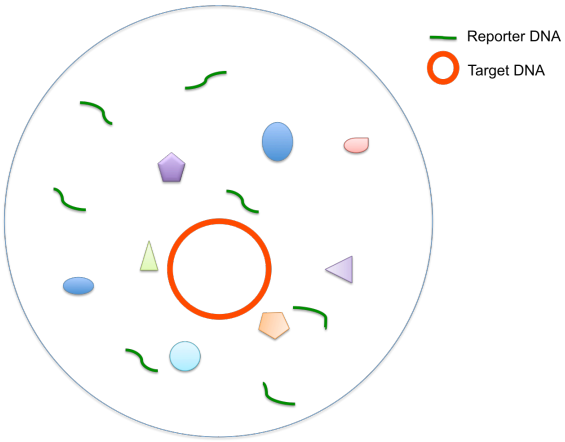


Figure 3.1 In-vitro two-hybrid-based detection strategy. (a) One-hybrid transcription regulation mechanism for DNA (encoding the protein in red) detection using DNA-protein interaction. Once the DNA binding domain (BD) binds to the upstream activating sequence (UAS) adjacent to the promoter (*psp*A) on the reporter gene, it activates the transcription initiation of the reporter gene by interacting with the promoter-bound E. coli RNA polymerase (RNAP) via the activation domain (AD), and hence triggers the translation of the fluorescent reporter gene (*venus gfp*). (b) Two-hybrid transcription regulation mechanism for DNA (coding for the protein in red) detection using protein (red)–protein (grey) interaction. An interacting protein pair X and Y are fused to BD and AD respectively, forming two hybrid proteins, which work together to activate the transcription. (c) Schematic of a drop containing the target DNA (in plasmid), reporter DNA (linear) and multicomponent IVPS kit. The components in IVPS kit are represented by the irregular-shaped symbols in the picture.

3.2.2 Target encapsulation and droplet IVPS

Figure 3.2a illustrates the overall operations of the microfluidic drop-based detection experiment. The mixture of the ice-cold IVPS kit and target DNA was introduced into the aqueous channel of the 10 μm drop-maker at a flow rate of 60 μl/h, and the fluorinated carrier oil Novec HFE-7500 with 1.8% (w/w) EA-surfactant was infused into the oil channel at a flow rate of 250 μl/h. Monodispersed drops of ~0.5 pL containing the IVPS-DNA mixture were formed at the flow-focusing junction (Length: 10 μm, Width: 10 μm, Height: 10 μm) of the microfluidic drop-maker (Figure 3.2b) and collected into a 1 mL BD Luer-Lok™ syringe (BD Medical) for off-chip incubation. The number of the DNA molecules encapsulated in one droplet follows the Poisson distribution

$$P(X = k) = \frac{e^{-\lambda}}{k!}\lambda^k$$

where P is the probability of having k DNA molecules in a drop when the average DNA molecule number per droplet is λ. The drops with target DNA inside will be containing

the fluorescent protein venus GFP after a series of cell-free transcription and translation reactions during the 37 °C incubation. We incubated the emulsion in dark for ~6 hours to allow for sufficient reactions in drops, after which the emulsion was re-injected into the 10 µm microfluidic detector for fluorescence signal detection (Figure 3.2b). The emulsion was flowed into the detector at 20 µl/h and spaced out with the fluorinated carrier oil Novec HFE-7500 (200 µl/h) at the arrow-shaped junction before entering the detection region where the single-file lined drops were scanned by the 488 nm laser (Newport) for analysis.

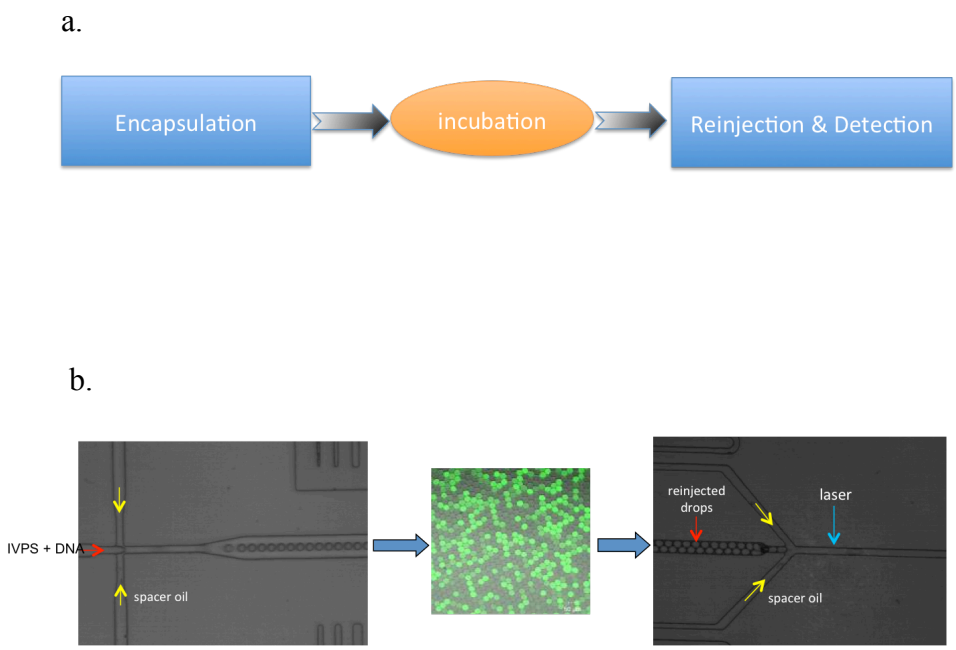


Figure 3.2 Overview of the DNA detection with drop-based microfluidics. (a) Schematic of the experimental steps for the microfluidic drop-based DNA detection. Off-chip operation is denoted in orange. (b) Representative pictures of the post-incubated emulsion (middle, overlay of the fluorescence image and bright field image), drop-maker (left) and detector (right). Emulsion flows from left to right in the detector. Channel depth: 10 µm. Narrowest channel width: 10 µm.

3.2.3 Detection of the single DNA molecules

- **DNA detection via protein-DNA interaction**

For its small size (~20 amino acids) and high binding affinity (K_d~pM) [57] to the consensus DNA binding site, we chose phage λ protein Cro as the target protein to replace the BD of PspF for the demonstration of the assay. As illustrated in Figure 3.1a, we were to detect the DNA *cro-ad* that encodes the protein Cro-AD at the single molecule level in the droplet, based on the interaction of Cro and the consensus binding site on the reporter gene construct.

To characterize the drop-based detection system, we first measured the fluorescence from the "empty" drops, in which no target DNA was introduced. Fluorescence intensity time traces data were recorded and plotted for analysis, as in Figure 3.3a. Each peak above the baseline level corresponds to a fluorescent droplet trajectory captured by the detector when it was transiting the detection window, with the peak width indicating the drop size information. The frequency of the peaks corresponds to the occurrence of the fluorescent drops at a given flow rate. In the experiment with no target DNA in the drops, there was a moderate fluorescence level observed in every droplet, as represented in a 10 ms snapshot of the fluorescence time traces (Figure 3.3a). This background fluorescence indicates the weak promoter *psp*A controlled expression level of the reporter fluorescent protein in the absence of the activator. Its consistent intensity among drops can serve as the droplet indicator and thus save the extra step of incorporating indicator fluorescent dyes in the droplet as used in some drop experiments

[48,55,58]. When an average of 10 DNA molecules ($\lambda = 10$) were encapsulated into the drops, a similar uniform fluorescence level over the majority of the drops was observed in the time traces plot, with a 20 fold increase in amplitude above the background obtained at $\lambda = 0$. And a close examination on the histogram of the drop fluorescence at $\lambda = 10$ confirms a small peak at the background fluorescence intensity (Figure S.1b), indicating the existence of a small population of the "empty" drops, which agrees well with the Poisson distribution ($P_{(X=0,\ \lambda=10)} = 4.54E\text{-}5$). A major difference arises in the single-molecule detection experiments, for example, when the target DNA was diluted to a concentration of $\lambda = 1$. The fluorescence intensity is no longer predominated by a single signature (Figure 3.3c). It is more clearly displayed in the histogram of the drop fluorescence (Figure 3.3d) where a significant portion of drops (~0.359) are "empty", in excellent agreement with Poisson distribution ($P_{(X=0,\ \lambda=1)} = 0.37$), and thus demonstrates the successful detection of the single-occupancy drops. Experiments with more diluted DNA ($\lambda < 1$, data not shown) further verify the system's capability in detecting single *cro-ad* DNA molecules in drops. The fluorescence intensity from the single-DNA drops can be obtained from the pivot point (most probable occupancy event) in the drop fluorescence distribution curve ($\lambda \leq 1$).

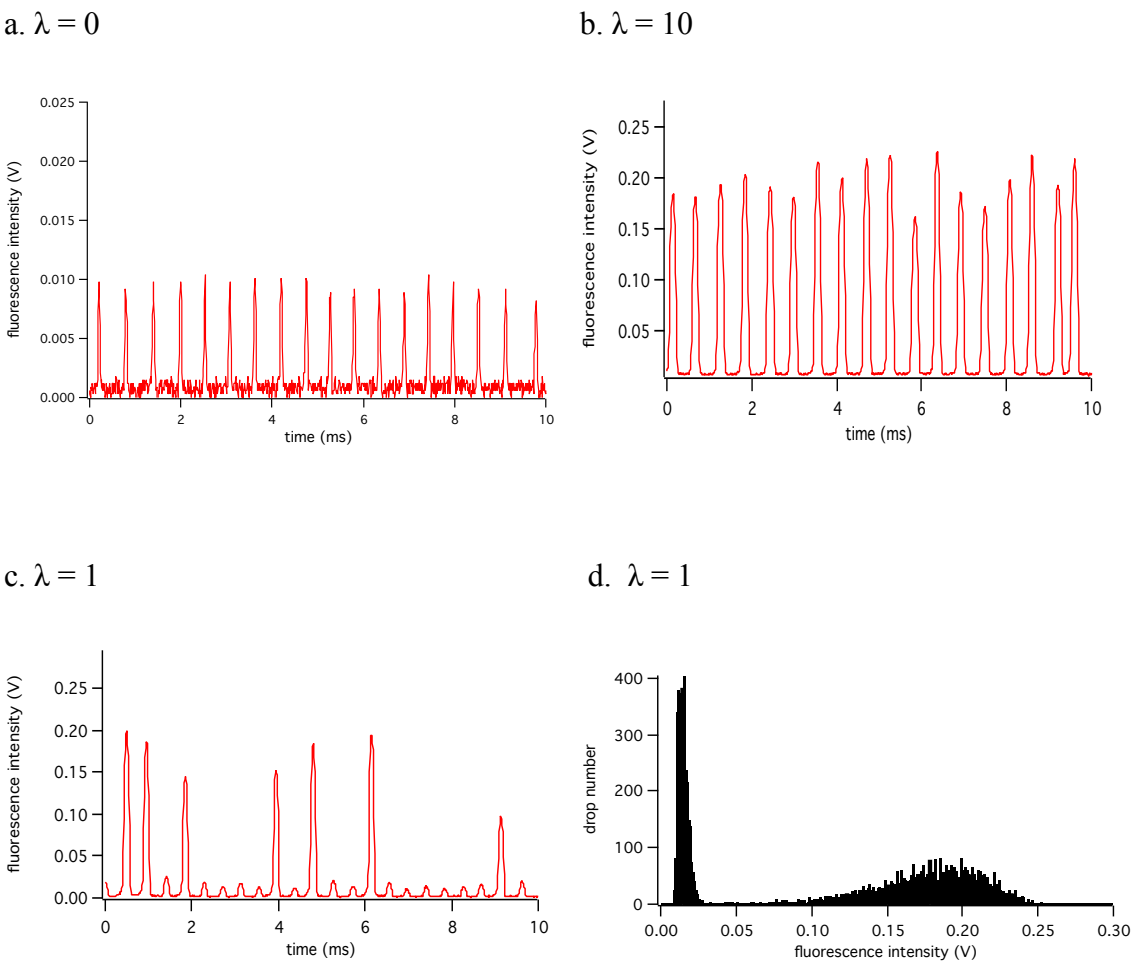


Figure 3.3 Detection of the DNA *cro-ad* in drops. Examples of the fluorescence intensity time traces at different *cro-ad* concentrations: (a) λ = 0, (b) λ = 10 and (c) λ = 1. Every peak above the baseline corresponds to a droplet. The height and width of the peak provide the information on the fluorescence and the size of the drop. The frequency of the peaks can be adjusted by the flow rates of the re-injected emulsion and spacer oil. (d) Drop fluorescence distribution obtained from the experiment at λ = 1. Drops were identified from the time traces data with the peak analysis algorithm executed on the FPGA.

As demonstrated above, for protein Cro-AD that has a high binding affinity to the consensus sequence on the reporter gene construct, our system can detect its DNA at the

single molecule level unambiguously. To test the system's tolerance to weaker binding affinities, we performed a series of detection experiments on the targets that encode zinc finger proteins as the synthesized BD. Zinc finger is another small (23-28 amino acids) DNA-binding protein that is widely used in protein engineering as a preferred scaffold for creating customized DNA-binding domains with improved binding affinity and specificity [51,52,59]. It binds to the DNA recognition site through multiple finger-like motifs coordinated by zinc ions. Via making changes to the original fingers or adding more fingers, higher affinity or specificity can be obtained in the zinc finger variants. For example, the reported binding affinity K_d of the wild type zinc finger Zif268 (Zif) to its wild type binding sequence is in the range of 10-500 pM [60,61], much lower than that of Cro to its consensus sequence. Correspondingly, our experimental results with *zif-ad* as the target DNA showed no discernible difference in fluorescence intensity between the "empty" drops and the drops containing *zif-ad* at $\lambda = 1$ (Figure 3.4a), suggesting that the transcription of the reporter gene was not effectively activated, due to the weak interaction between the BD Zif and the cognate sequence. Whereas a variant Zif//NRE screened by J.-S. Kim and C. O. Pabo [61] is claimed to have a much higher binding affinity (<1 pM) to its binding site. In our detection experiments with *zn-ad* (encoding Zif//NRE-AD), we were able to observe the two populations of "empty" and occupied drops at $\lambda = 1$ (Figure 3.4b), indicating the existence of the fluorescence from the target-containing drops with successful activation on transcription, hence suggests a stronger interaction of Zif//NRE to its binding site as reported. And the peak point in the occupied population corresponds to the most probable drops whose fluorescence, in the case of $\lambda = 1$, indicates the fluorescence from the single DNA-bearing drops. In addition to the

successful detection of the target DNA at the single molecule level as discussed above, the stronger protein-DNA interaction from Zif//NRE-UAS_{zn} compared to Zif-UAS_z suggested from our experiments is in good accordance with the measured binding affinities for those two protein-DNA interaction pairs, which verifies the positive correlation between the gene expression level and the interaction strength of its transcription activator and UAS in our drop IVPS-based two-hybrid system. It supports the similar finding in bacteria-two-hybrid system about the correlation between the transcription activity and the interaction strength of the activator protein pair [59], and enables the system to indicate the molecular interactions involved in the regulation mechanism quantitatively in addition to the DNA detection.

Aforementioned detection experiments were all performed on the single-type target in an IVPS droplet. In fact, our system could also detect the target in a multiple-DNA cocktail, as exemplified in Figure 3.4c, where the target DNA *cro-ad* and control *zif-ad* (as to the consensus sequence of Cro) were mixed at 1:10 with a total λ of 1. The population containing the target was identified as the small hump in the drop fluorescence histogram. It thus encouraged us to try detecting more types of DNA based on more complicated molecular interactions.

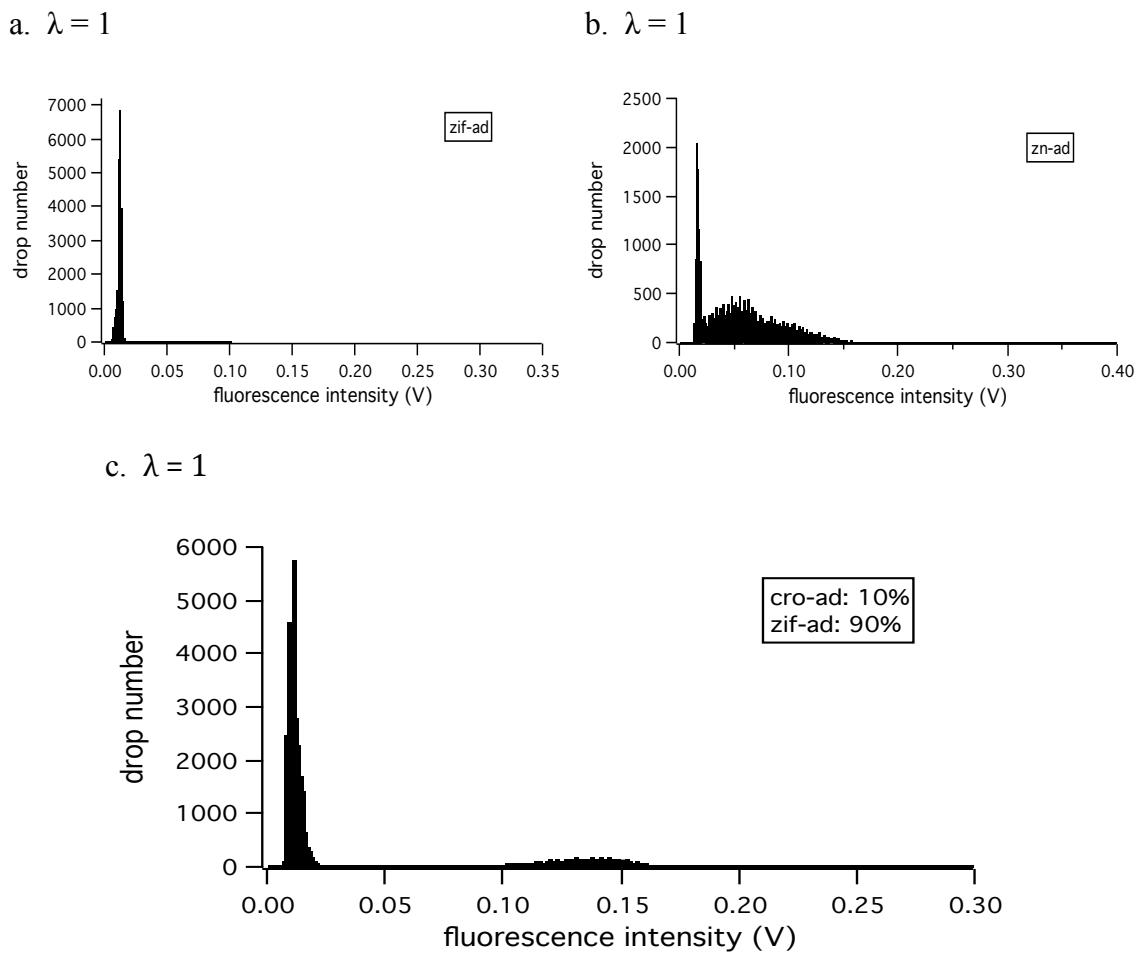


Figure 3.4 Detection of the single molecule DNA via weak protein-DNA interactions. Drop fluorescence intensity histograms obtained from the experiments with different target DNA templates at λ = 1: (a) *zif-ad*, which encodes a protein with the wild type Zif268 as the DNA binding domain. (b) *zn-ad*, which encodes a protein with the variant Zif//NRE as the DNA binding domain. PMT gain was raised to 0.5 for clarity. (c) Mixture of *zif-ad* and *cro-ad* (λ_{total} = 1) with the ratio of 9:1.

- **DNA detection via protein-protein interaction**

So far, we have successfully detected the single target DNA based on different protein-DNA interactions in the transcription regulating mechanism. However, not all the

proteins have a fair binding affinity with UAS to be the activator in the protein-DNA interaction-based transcription regulation system. To expand the application of the assay, we challenged our system in detecting the DNA single molecules based on protein-protein interactions (Figure 3.1b), where the activator BD-X synthesized from the target DNA in the drop works in concert with a partner protein Y-AD to initiate the transcription of the reporter gene. If there is enough strength of the specific interaction between X domain and Y domain, the protein pair will function as one activator to initiate the expression of the reporter protein, thus leading to the detectable fluorescence signals from the droplet. As a demonstration, we used the interaction between human estrogen receptor α (ER) and nuclear receptor co-activator 1 (NCOA1) under the modulation of the small molecule 17β-estradiol (E2) to detect the target DNA bd-ncoa1 at the single molecule level.

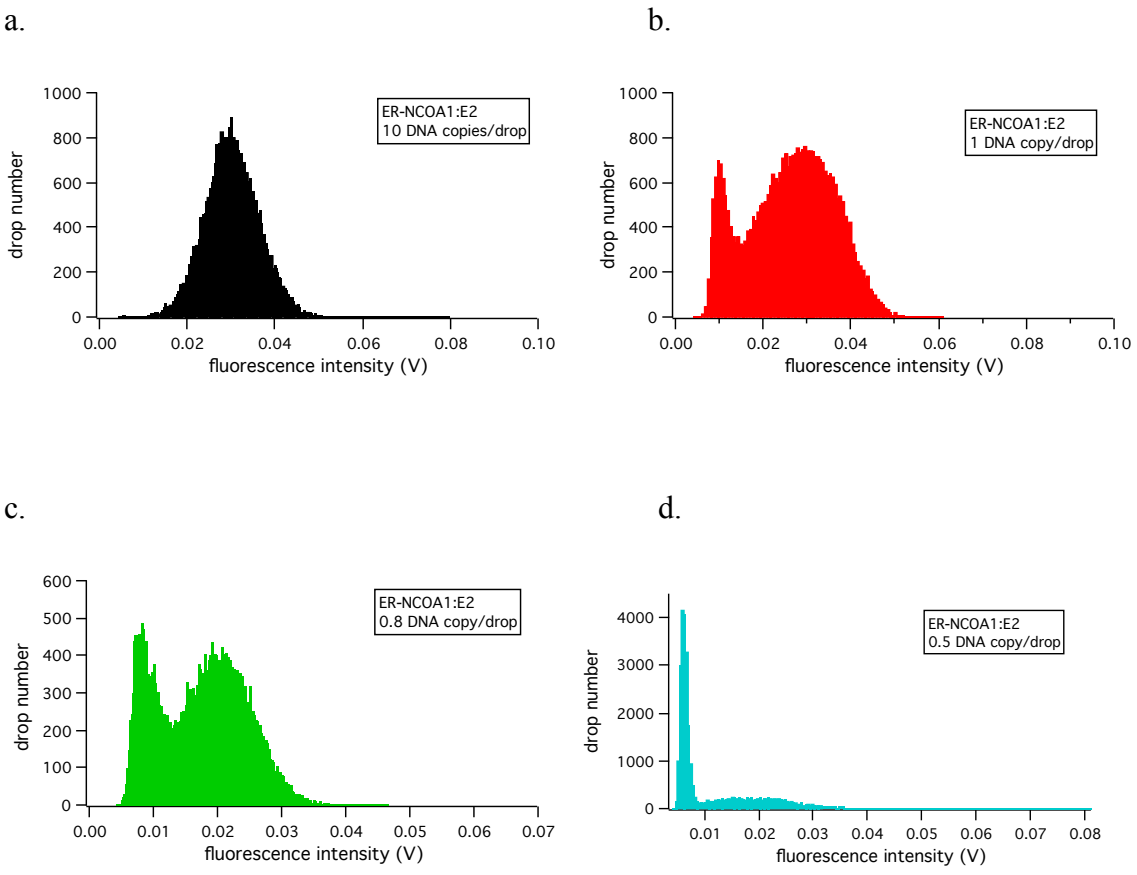
The partner protein ER-AD was synthesized in IVPS reaction along with the target protein BD-NCOA1 to trigger the synthesis of the reporter protein venus GFP in the droplet. As discussed above, the amount of the DNA molecules for the partner protein and reporter protein as well as the modulator molecules were optimized for the maximal production of the reporter protein within the limited resources in IVPS. Target DNA was co-encapsulated with the partner DNA ($\lambda = 10$), E2 (1 μM) and IVPS kit into the 10 μm-diameter droplet before the reaction-inducing incubation off-chip. We measured the fluorescence of the re-injected droplets at ~2 kHz at a series of target DNA concentrations (Figure 3.5).

At $\lambda = 10$ (Figure 3.5a), the majority of the drops showed a fluorescence signal above the background level, indicating a predominant DNA-occupancy among the drops.

The presence of the small population at the background fluorescence level agreed well with the Poisson distribution. In the experiments of $\lambda \leq 1$ (Figure 3.5b-3.5f), the two populations of the drops became increasingly discernible in the drop fluorescence histogram, with the lower fluorescence drops gradually outnumbering the higher fluorescence ones as λ decreased, indicating the increase in "empty" drop number and single-occupancy occurrence. The consistence of the experimental drop distribution with Poisson distribution verifies the detection of the single *bd-ncoa1* molecule. It not only demonstrates the robustness and sensitivity of our detecting system in the ability to incorporate a complex multicomponent in-vitro protein synthesis, but also expands the range of the detectable DNA by introducing a partner protein to allow for a variety of molecular interactions to be utilized for the detection. Here we used BD-NCOA1 as the target protein; similarly ER-AD can be set as the target protein as well if the optimized amount of *bd-ncoa1* is fixed. Based on the same principle, proteins that have interactions with other biomolecules, for example, RNA, can also be the target protein for the detection of its DNA, where RNA can be tailored as a bridge between the activator pair to initiate the expression of the reporter protein.

In the *bd-ncoa1* detection experiments, we noticed a subtle λ-dependent ($\lambda \leq 1$) left-shift trend of the fluorescence intensity in both the single-occupancy and the "empty" drops in the drop fluorescence histogram. In other words, at the single molecule level, as the target DNA concentration decreased, the fluorescence level from "empty" drops and the drops containing the single target DNA molecule both decreased in a subtle way. However, no similar changes were observed in the $\lambda > 1$ experiments. The constant fluorescence level in the $\lambda > 1$ experiments can be explained as the IVPS reached its

plateau for the synthesis reactions incorporated and therefore any increase in the target DNA amount would not bring in the increase in the production of the reporter protein. But interpretation to the intensity shift at the single molecule level is not that straightforward. Since the fluorescence was from the fluorescent protein that has a large molecular weight compared to the small fluorophore molecules, the chances of the fluorophore exchange between drops due to the concentration gradient observed in experiments with fluorescent dyes [56,62,63] are slim. Nevertheless, this subtle shift of both signal and background levels in the diluted DNA experiments does not affect our ability in detecting the single DNA molecule.



e. f.

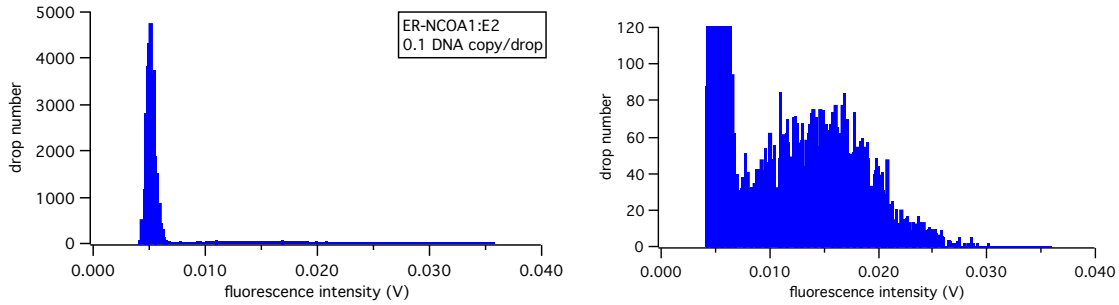


Figure 3.5 Detection of the DNA *bd-ncoa1* in drops. Drop fluorescence intensity histograms at different *bd-ncoa1* concentrations: (a) $\lambda = 10$, (b) $\lambda = 1$, (c) $\lambda = 0.8$, (d) $\lambda = 0.5$, (e) $\lambda = 0.1$. (f) is a zoom-in of (e) on the low y-value region for a clear view on the high-intensity population. The gene encoding the partner protein ER-AD was kept at $\lambda = 10$, and E2 was kept at 1 uM in all the *bd-ncoa1* detecting experiments.

3.3 Conclusion

For the first time we demonstrate the capability of drop-based microfluidics in label-free, PCR-free high-throughput detection of specific DNA at the single molecule level. The system detects target DNA through monitoring the interaction of its protein with other biomolecules in a pico-liter droplet. Since the detection is based on its inherent physiological functions, there are no concerns of the potential damage or block to the properties of interest due to the labeling of the target DNA. On the contrary, the detection strategy we proposed here can shed light on the molecular interactions monitored in the droplet alongside the DNA sensing, as discussed above. Benefited from the development in the in-vitro protein synthesis techniques, the assays in our detecting experiments are robust and highly reproducible, enabling the reliable detection of DNA in an intricate dynamic biological environment without being significantly affected by any perturbation

from the system's complexity. It therefore provides more access to studying the fundamental biology that requires the isolation of the target gene from its ensemble [19]. Moreover, the advantageous features in the reaction volume and throughput of the drop-based microfluidic platform confer this detection technique great sensitivity (single molecule level) and ultrahigh-throughput (kHz) simultaneously, which especially benefits the screening and sorting of the target from a large library as in the protein engineering and drug screening field (unpublished data). For example, we could use this system to detect p53 mutations in the cancer cell. Most of its mutations happen in the BD region that would consequently cause a failure in recognizing the DNA binding sequence and thus the inability in activating the DNA repair pathway to eliminate mutagenesis in cells.

Chapter 4

SOS Pathway-Based Fluorescence-Activated Drop Sorting System for Ultrahigh-Throughput Screening of Restriction Endonucleases

4.1 Introduction

Restriction Endonuclease (RE) is an enzyme that can recognize a particular DNA sequence and cleave DNA specifically into fragments with double-stranded breaks (DSBs). Its sequence-specific cleaving ability has enabled it a highly useful gene-editing tool since its first isolation and characterization in 1970s [64-66]. Apart from creating recombinant DNA in conjunction with DNA ligase, RE also has applications in gene analysis, for example, characterizing gene mutations through analysis on the cleaved fragment length [67,68]. With the increase in its availability and specificity, RE is now being extensively studied as a promising agent for targeted gene disruption in the gene therapy for various virus diseases, such as HIV and HPV [69-73].

The conventional approach of searching for new REs is based on the restriction-modification (R-M) system where potential RE genes are predicted through data mining techniques from established sequence database according to their cognate methyltransferase genes which are recognizable by bioinformatics methods because of the conserved motif elements, and in most cases, reside close to RE genes. The predicted putative RE genes are therefore not available until being verified through biochemical

characterization [74]. The traditional labor-intensive biochemical characterization method, however, is having a hard time in catching up with the explosive increase in putative RE database that benefited from the development in DNA sequencing technology and real time methyltransferase detection technique [75,76]. Moreover, not all the RE genes originate from the R-M system where a methyltransferase gene is adjacent, for example, Pad and Pmel. Therefore, more efficient RE generating methods are in great demand.

In recent years, artificial REs that consist of a DNA binding domain and a cleaving domain, such as zinc finger nucleases (ZFNs) [77-83] and transcription activator-like effectors (TALEs) [84-91], have attracted great attention as possible substitutes for natural REs. Their modular structure provides a feasible means to generate more REs through rational designs on each modular domain. However, the context-dependence on neighboring residues and target DNA puts an obstacle in obtaining high binding-specificity ZFNs [92,93]. As for TALEs, the relatively large size appeared to be a challenge in its assembly [85,89,91]. The newly discovered CRISPR/Cas9 system [94-106], recognizing the target DNA sequence through complementarity of base-pairs with its guide RNA molecules, seems like a most promising gene editing tool with the convenience, in theory, of targeting any specific DNA sequence with a high flexibility and simplicity in design and assembly. Yet more experimental data on the successful application of this system to the target are needed to characterize and validate the proposed systems.

To facilitate the verification and therefore the development of the advantageous natural or artificial REs, we present a novel strategy developed for ultrahigh-throughput

RE gene screening using fluorescence-activated droplet sorting (FADS) technique based on the prokaryotic cell SOS pathway. With this strategy we demonstrated the successful enrichment on the target RE gene XbaI from a model library consisting of XbaI and its truncated mutant gene ΔXbaI with a series of library sizes. The kilo-Hertz interrogating rate enables this technique promising in efficiently screening RE genes from, for example, the putative RE library predicted from the R-M systems, or from the CRISPR/Cas systems and other artificial RE libraries that are being studied. Moreover, it can also be applied to large-sized metagenomic libraries constructed from environment samples that have a high possibility of containing new REs with novel features such as better stability, high pH or high temperature tolerance.

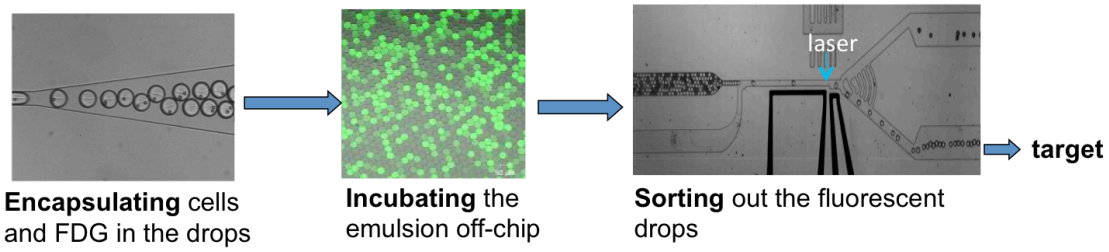
4.2 Results and Discussion

4.2.1 RE screening scheme

Figure 4.1a illustrates the overall operating steps for RE screening on the drop-based microfluidic platform. The Escherichia coli (E. coli) cells are encapsulated in the medium together with fluorogenic substrate Fluorescein-Di-β-D-Galactopyranoside (FDG) and Isopropyl β-D-1-thiogalactopyranoside (IPTG) to form a droplet of 23 μm in diameter. Following the off-chip incubation at 37 °C for a few hours, the emulsion is then injected back into a microfluidic chip for fluorescence signal detection and target sorting. The drops containing the cell with RE gene will be fluorescent when they pass by the detection window illuminated with excitation laser. The biology mechanism underlying the fluorescence is described in Figure 4.1b. Briefly, RE is expressed under T7 promoter on the plasmid vector pTXB1 upon the addition of IPTG that initiates the transcription

through the lac operon. The attack of REs leads to the double stranded breaks (DSBs) on the genome of the host cell, which activates a series of gene expressions in DNA repair pathways, including the overexpression of β-galactosidase (β-gal) whose gene is fused to the downstream of the damage sensing promoter dinD1 in the SOS pathway. The excessively expressed β-gal helps hydrolyze FDG and thus produces the fluorescent molecule fluorescein.

a.



b.

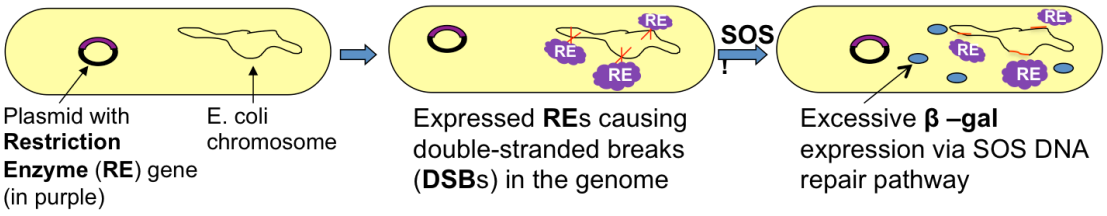


Figure 4.1 Restriction Enzyme (RE) screening scheme. (a) Overview of the operation steps for RE screening on the microfluidic platform, which includes encapsulation of the host cells and fluorogenic substrate, off-chip incubation and detection and sorting on the re-injected emulsion. (b) Biological mechanism that enables the RE gene detectable on the microfluidic platform.

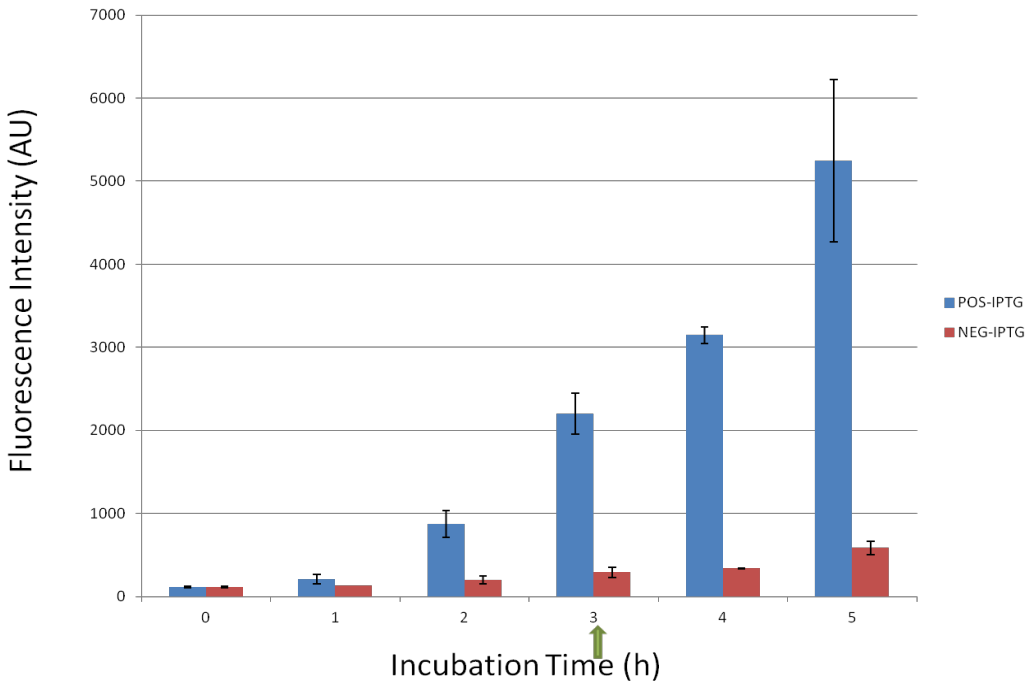
4.2.2 Verification of the RE screening scheme in bulk and in drops

This SOS-pathway based RE gene detection strategy was first verified in the bulk experiment where the "+" cells (ER 2745-pLysY-pTXB1_XbaI), after sonication exaction of the indicator protein β-gal, were loaded into the microtitre plate filled with FDG and IPTG. The time-course measurement on the green fluorescence intensity was performed in a 96-well microplate reader at 37 °C for 5 hours. The fluorescence intensity kept increasing with time whereas the control measurements on the "-" cells (ER 2745-pLysY-pTXB1_ΔXbaI) with the truncated RE gene showed no significant changes in the intensity (Figure 4.2a).

The fluorescence detection of the drops on the microfluidic platform showed similar results (Figure 4.2b), where drops containing "+" cells gave a distinguishable fluorescence signal compared to the empty drops or "-" cell-loaded drops. The difference in the signal intensity between the "+" cell drops and "-" cell drops was still discernible in the experiments where the two types of cells were mixed up and co-encapsulated in the drops. Noted that in the droplet detection results, the high background fluorescence from the "leaky" control of the lac operon in this host cell strain was displayed more clearly than in the bulk measurements, because the fluorescence histogram of individual drops was able to show the distribution of the "empty" drops (no cells included) when the cell density was diluted to be less than one copy per droplet on average, which was buried in the bulk measurements. The close locations between the two peaks in the distribution curves from the experiments with "+" cells incorporated indicated that the signal to noise ratio in this system was not ideal, which could be further improved by refining the host cell construct for a lower background expression level of the indicator protein. For

example, another derivative strain ER 2746 from NEB possesses similar activities but with a lower leaky expression of β-gal. It therefore can be used as a substitute if we don't aim for a higher than 37 °C reacting temperature. The improvement on the host cell system can also be considered on the selection of other DNA damage promoters or protein indicators.

a.



b.

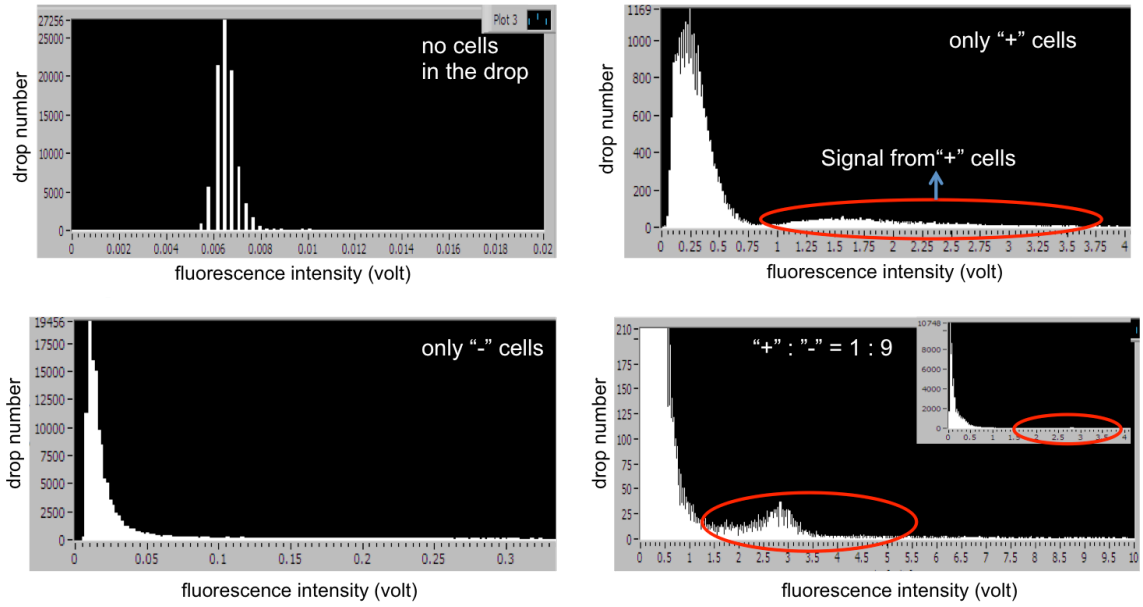


Figure 4.2 Fluorescence measurements on the RE-containing cells and the control cells. (a) Time-sequence fluorescence measurements on the XbaI-containing ("+") cells and ΔXbaI-containing ("-") cells in the plate-reader where fluorogenic substrate Fluorescein-Di-β-D-Galactopyranoside (FDG) and inducer Isopropyl β-D-1-thiogalactopyranoside (IPTG) were included in the well. The chamber was maintained at 37 °C during the measurements. The green arrow indicates the off-chip incubation time in the droplet experiments, which is a trade-off between the fluorescence induction and FDG's stability. (b) Fluorescence measurements on the drops with no cells, pure "+" cells, pure "-" cells, and a mixture of "+" and "-" cells after 3 hours' incubation at 37 °C in dark. The average cell number per drop was kept at 0.3 in all the experiments. Red circles in the plots where "+" cells were involved indicate the population of cell-containing drops in the drop fluorescence distribution.

4.2.3 Single round of sorting on model library

Based on the detection results, we performed the RE gene sorting experiments on the model library consisting of "+" cells that were transformed with XbaI gene, and "-"

cells that were transformed with ΔXbaI gene. For libraries with initial size[1] of 1:10, 1:100 and 1:1000, the target XbaI genes were all successfully enriched in the sorted samples, as examined from the agarose gel electrophoresis results (Figure 4.3). XbaI and its mutant gene ΔXbaI were amplified through colony PCR from the sorted samples. The size difference between the two genes allows us to determine their relative amount by comparing the gel band brightness.

For the library with the initial size of 1:10, we got an enriched ratio of 1:4 in the sorted sample, which gave an enrichment of 2.5 folds. Similarly, the enrichment for the library with the initial size of 1:100 was 33.3, and the same enrichment value was obtained for the library with the initial size of 1:1000. The low enrichment for the 1:10 library might be attributed to the lower sorting gate, which allowed the entrance of the "-" cell drops that possessed a background fluorescence from the basal expression of β-gal in the host cell into the sorted channel. Therefore, the optimal enrichment level from the single round of sorting can be achieved through a fine tune of the sorting gate and input cell density.[2]

The similar enrichment levels in the library of 1:100 and 1:1000 obtained from our experiments are consistent with the theoretical calculation given by Y. Zheng [107], who provided a theoretical enrichment prediction based on Poisson distribution for his droplet directed evolution experiment, where the RE genes were sorted by the adaptor-specific PCR (Figure S.2). However, the calculation only takes account of the co-encapsulation effect from Poisson distribution as the false positive contributor. But in our

[1] Initial size means the ratio of the number of "+" cells to that of "-" cells when the cells were mixed and encapsulated in the drops.

[2] We used $\lambda = 0.3$ in all of our RE screening experiments (see λ meaning in Chapter 3).

droplet experiments, the fluorescence from the "-" cell drops might also contribute to the false positive hits in the sorted sample, especially in the low-gate sorting. Apart from that, the cells were suspected to keep dividing in the drops during the signal-inducing off-chip incubation, and the different dividing rates for "+" and "-" cells due to different DNA inserts might further explain the disparity between the predicted enrichment values and our experimental values.

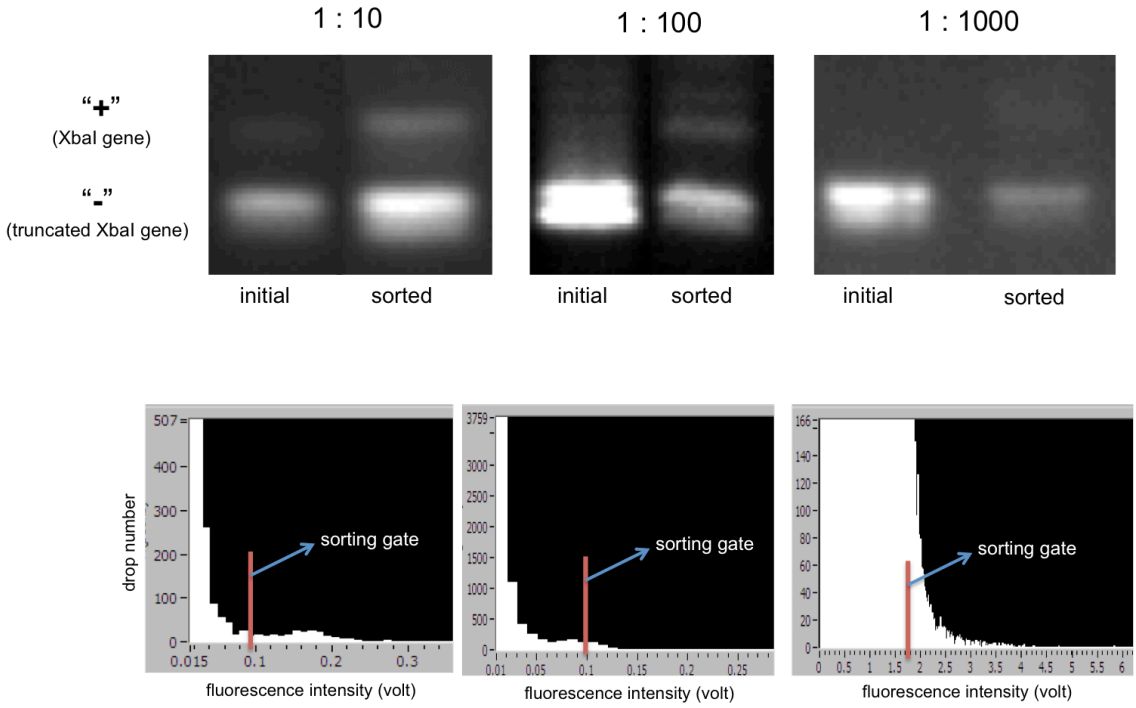


Figure 4.3 Single-round sorting results for different initial library sizes: 1:10, 1:100 and 1:1000. The sorting gate selected in each individual screening is indicated in the red line. The quantification of the enrichment was performed with ImageJ (NIH).

4.3 Conclusion

As discussed above, through the successful enrichment of the target RE genes from a constructed model library, we proved the principle of the SOS pathway-based microfluidic sorting strategy for RE gene screening. By optimizing the system through, for example, carefully designing the average cell number per droplet and specifically selecting the sorting gate for each round, it is promising to find the hit RE genes efficiently with this approach in the real libraries after multiple rounds of sorting. In fact, the reports on successfully selecting horseradish peroxidase variants with enhanced catalytic activity [14] and identifying xylose isomerase genes that contributes to high xylose consumption from the genomic library [108] through multiple rounds of sorting on a similar drop-based microfluidic platform raise much hope for the practicality of this approach.

It is the first time the prokaryotic DNA repair mechanism has been integrated into ultrahigh-throughput microfluidic system to realize the signal transduction from the target genotype to the detectable phenotype with efficient signal detection and processing, which is especially beneficial for large library screening that usually requires multiple rounds of sorting. In our screening scheme, the overexpression of the reporter gene β-gal was induced by the DSBs on the host genome as a result of the attack by REs. Therefore, this system, after the necessary adjustment, can also be used to screen for other enzymes or molecules that can cause the DSBs in the host genome. Furthermore, by incorporating the in vitro transcription and translation reagents into the drop-based microfluidic sorting system, we can even expand our screening scope from being based on this particular

dinD1 repair pathway to other intrinsic cellular pathways, which will be of great significance in both protein engineering and biomedical applications.

Chapter 5

Microfluidic Pico-injection Technique for Controllable Fabrication of Less-Invasive Alginate Gel Beads as the Three-Dimensional Extracellular Matrix for In-Vitro Stem Cell Niche

5.1 Introduction

Stem cell niche, a highly specialized and dynamic microenvironment where stem cells reside, has been recognized for its essential regulating role through its components in stem cells' survival, self-renewal, differentiation and communication with the surrounding tissue [109-114]. As its major component, extracellular matrix (ECM) not only provides necessary mechanical and adhesive support to the cells, but also regulates various cell behaviors through facilitating the transmission of biochemical cues [112,115-119]. Therefore, systematic studies on the interactions between the stem cell and its niche are of great significance in understanding the mechanisms underlying the intricate stem cell behaviors [120-125] and advancing the development of tissue engineering techniques for effective cell therapies [126-132].

To apply reductionist approaches on stem cell niche studies, researchers have been extensively investigating on the methodology for synthesizing in-vitro ECM scaffolds that imitate the physiological environment [133-140]. Among many promising

ECM materials, alginate hydrogel has been widely studied because of its good biocompatibility as a natural ECM secreted from some bacteria and algae, and unique structural properties that are amenable to various chemical functionalization [141-152]. Alginate is the linear polysaccharide composed of repeating monomer units: β-D-mannuronic acid (M) and α-L-guluronic acid (G). Exposure to the divalent cation such as Ca^{2+} initiates the gelation process that crosslinks G blocks and forms the 3-D hydrated gel network. The considerable amount of hydroxyl and carboxyl sites on the polymer chain enables convenient chemical modifications to the hydrogel for desirable, for example, mechanical property and biodegradability. For the objective of stem cell niche studies where the cells need to be cultured in the scaffold for enough length of time for observations, for example, on the proliferation and differentiation behaviors, it is ideal to minimize the size of the synthesized ECMs to avoid the diffusion barrier for nutrients, oxygen and biochemical cues. However, the previously reported alginate microgel synthesis approaches are either limited by the large pore size and potential adverse effect on the cell behaviors due to the change of pH condition as in the internal gelation strategies [153-160], or have a low productivity and controllability as in some external gelation methods [161-168].

To provide a high-throughput, controllable method to synthesize minimum-invasive alginate microgel scaffolds, we develop a microfluidic two-step strategy based on pico-injector technique, where drop-making process and gelation process are separated apart. Unlike the co-flow technique utilized in typical microfluidic alginate gel bead synthesis approaches, the micron-sized pico-injector nozzle effectively limits the contact area of the precursor solution and the cross-linker fluid, and therefore greatly

suppresses the occurring of the clogging due to the fast crosslinking in the ionic gelation process, thus making the synthesis of alginate gel beads more controllable, reliable and efficient on the microfluidic high-throughput platform. Moreover, no additional chemicals or pH condition changes are involved in our two-step approach; therefore, the synthesized alginate gel beads are expected to be less invasive in the long-term cell culture for behavior studies on these stem cells.

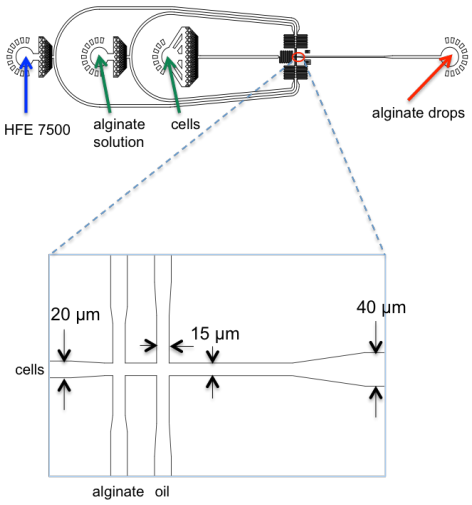
5.2 Results and Discussion

5.2.1 Two-step gel synthesis process

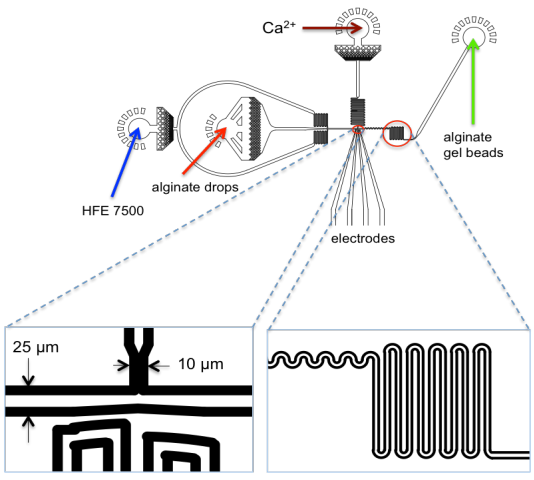
Figure 5.1 describes the two-step synthesis process, with the drop-maker (Figure 5.1a) producing aqueous alginate drops, followed by the pico-injector (Figure 5.1b) serving as the gelation platform. The narrow size of the injecting nozzle reduces the possibility of clogging at the interface of the precursor and cross-linker fluids, whereas the evenly spaced single-file flow of the alginate drops allows a tight control of the cross-linker addition for an improved uniformity on the gel bead structure. The incorporation of the sinusoidal design in the downstream of the T-junction enables an efficient mixing of the calcium ions and alginate polymers in the droplet from advection, and the subsequent serpentine channel allows for the sufficient gelation before the drops cream in the collection tube. Typical drop-making process is demonstrated in Figure 5.1c, where encapsulated mouse Mesenchymal stem cells (mMSCs) are highlighted with the red arrows. Figure 5.1d is the time-sequence snapshots from a pico-injecting process captured by the fast camera (Phantom V7.3; VisionResearch). The entire cross-linker-adding process was completed within 1.25 millisecond, which gives a kilo Hertz

production rate. By controlling the cell density infused into the drop-maker, we can control the cell number in each droplet. Through adjusting the flow rates in the pico-injecting process, we can control the ratio of the cross-linker and precursor in each gel bead to obtain controllable mechanical structure. And because of the clean gelation process, the pico-injector devices could usually be reused for multiple rounds, which greatly reduced the operation cost in time and money.

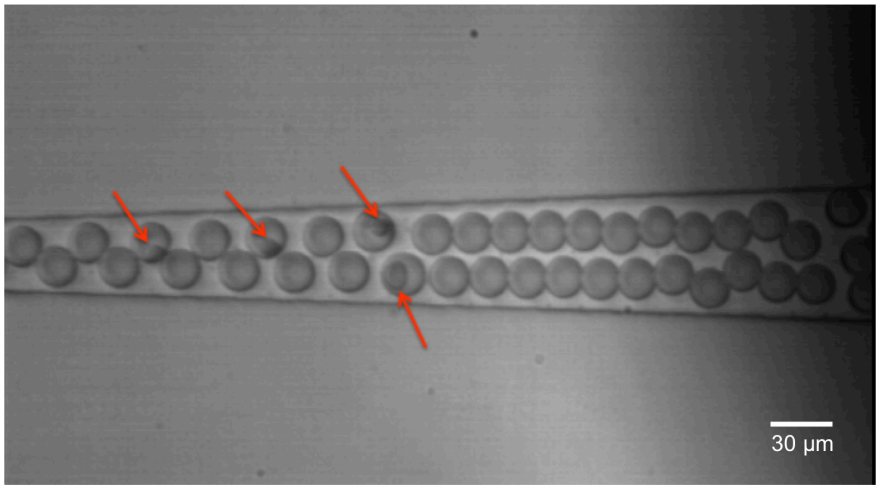
a.



b.



c.



d.

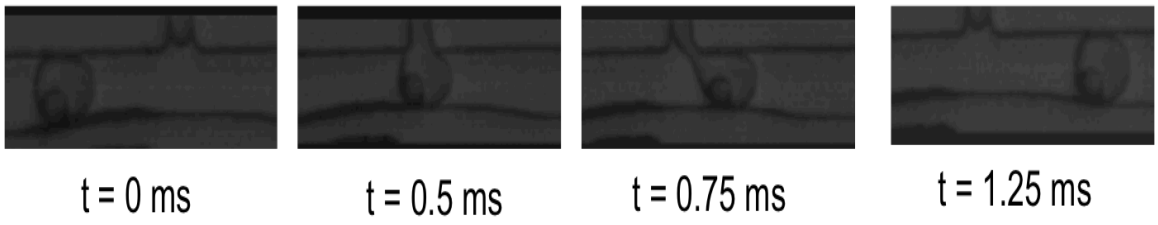
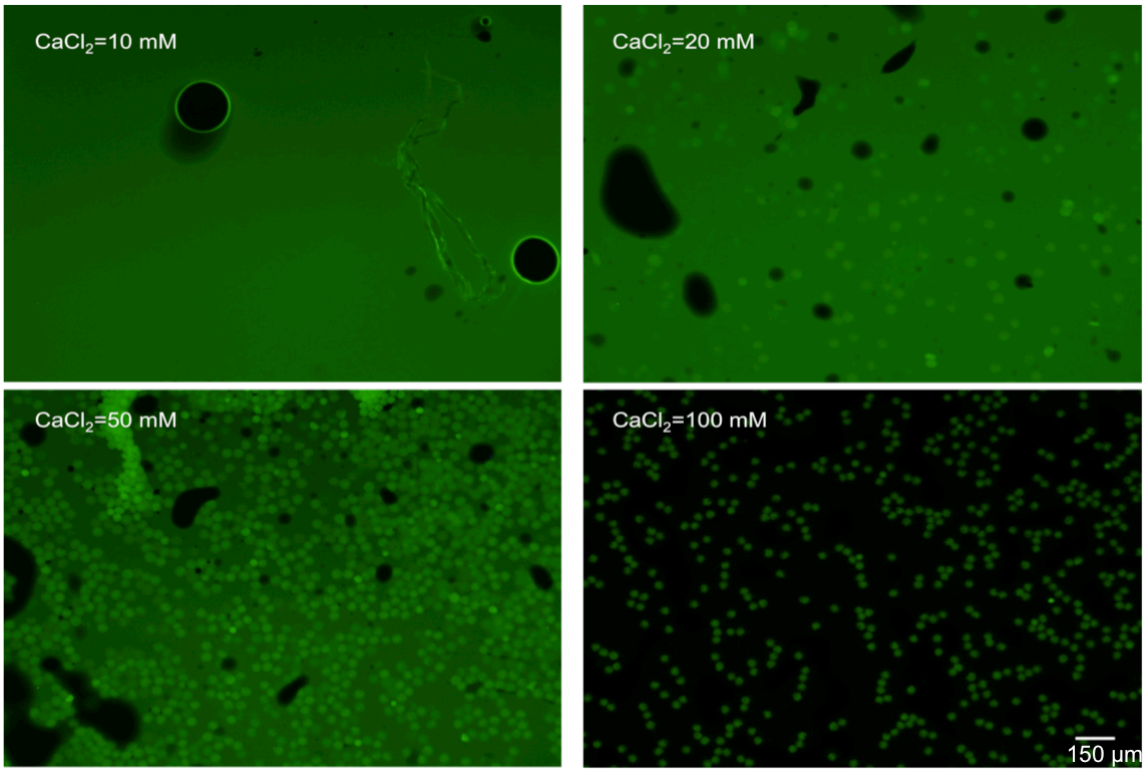


Figure 5.1 Two-step synthesis of alginate gel beads with microfluidic devices. (a) Schematic of the drop-maker device: Alginate solution and mouse Mesenchymal stem cells (mMSCs) were co-flowed into the aqueous channel from left to right. Fluorinated oil HFE 7500 with EA surfactant was flowed into the oil channel and sheered the aqueous fluid into ~25 µm drops at the junction downstream of the serpentine resistant channel. (b) Schematic of the pico-injector device: Alginate drops were injected into the inner channel and spaced out by the carrier oil HFE 7500 before flowing right to the T-shaped pico-injecting nozzle. The characteristic design downstream to the injecting junction (highlighted in the red circle) is intended for improved mixing and sufficient gelation on chip. (c) Snapshot of the movie of cell encapsulation into alginate drops. The encapsulated mMSCs were highlighted with the red arrows in the picture. (d) Time-sequence snapshots from a pico-injecting movie. In the representative snapshots, the cell-laden alginate drop was injected with calcium chloride from the perpendicular pico-injecting channel for gelation initiation. The fusion of the injecting cross-linker solution into the alginate aqueous drop was accomplished upon the activation of the electric field applied at the injecting area. The injection step was finished within 1.25 msec.

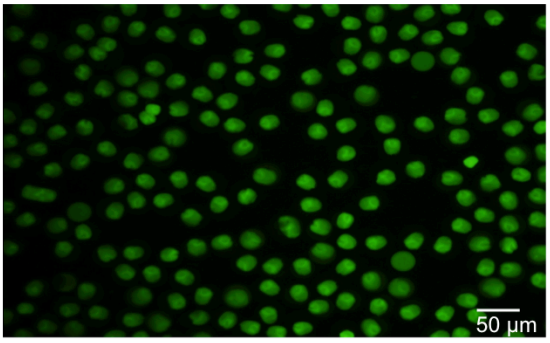
5.2.2 Gelling parameter optimization

With this two-step approach, we first synthesized empty alginate gel beads to obtain the optimum gelation conditions. We found that both the calcium concentration and its injected rate affected the gelation process. For the final concentration of 1% alginate, a majority of the stable gel beads that could keep the integrity after being washed out of the carrier oil into the cell medium were obtained only after the concentration of the calcium chloride was above 100 mM in the injecting channel (20 mM in the droplet), as indicated in Figure 5.2a. And the injecting rate of the calcium chloride had a great effect on the homogeneity of the meshwork in the gel bead. For a fixed calcium concentration, the increased injecting rate to the alginate drop resulted in an increased homogeneity in the spatial distribution of the crosslinking (Figure 5.2b). However, the increase in the injecting rate lead to the gel size increase, and in a confined channel, resulting in an anisotropic polar shape. For a trade-off between the internal homogeneity and the isotropic shape, we chose a calcium concentration of 20 mM in drops injected at a 20% rate (v/v) of the final droplet volume.

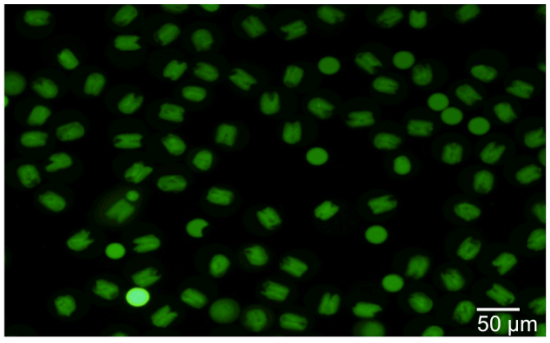
a.



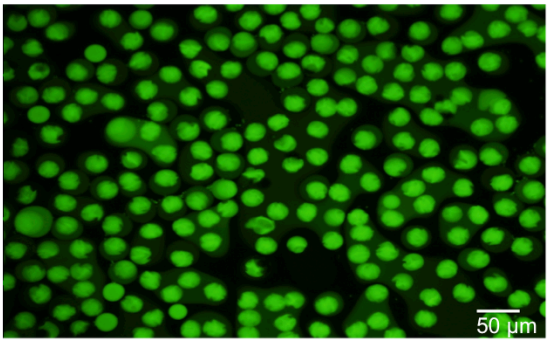
b.



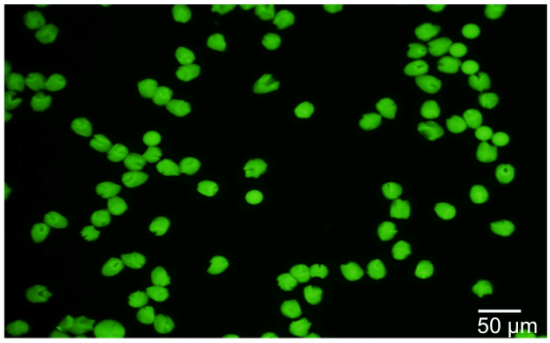
10 mM $CaCl_2$ 1:10 pico-injection



10 mM $CaCl_2$ 1:2 pico-injection



10 mM $CaCl_2$ 1:5 pico-injection



20 mM $CaCl_2$ 1:5 pico-injection

Figure 5.2 Optimization of the gelation conditions. (a) Fluorescence micrographs of gel beads formed from different Ca^{2+} concentrations. The displayed values are the concentrations in the pico-injecting channel. (b) Fluorescence micrographs of gel beads formed with different Ca^{2+} injecting rates. The ratios are the volume ratio of the injecting fluid to the injected alginate drop. $CaCl_2$ concentrations are the expected concentration in the injected drop. A fraction of alginate polymers were pre-functionalized with green fluorophore FITC for easy identification. The gel beads in (a) and (b) were both washed from the carrier oil into the cell culture medium before imaging.

5.2.3 Encapsulation of Mouse Mesenchymal stem cells in gel beads

We then incorporated mMSCs into the gel beads by encapsulating them in the alginate drops. For necessary cell adherence, alginate polymers were modified with the cell-binding peptide arginine-glycine-aspartate (RGD). For easy optical identification, a fraction of RGD peptides were also functionalized with green fluorophore Fluorescein isothiocyanate (FITC). Figure 5.3a (top left) shows the fluorescence image captured right after the gel beads were formed. The beads that contained the cells are indicated in red circles. The examination on the Trypan blue-stained cells in the phase contrast microscopic images showed more than 90% viability of the gel-encapsulated cells, as represented in the top right picture in Figure 5.3a. Similar cell encapsulation experiment was also performed with the red membrane dye PKH26-stained cells and the fluorescence images were taken right after the gel beads formation. As clearly displayed in Figure 5.3a (bottom), quite a number of cells were encapsulated in the ~40 μm alginate gel scaffold, with the existence of both single- and double-occupancy. The observation that cells tended to reside at the edge of the gel bead might be associated with non-uniform

crosslinking within the gel beads, which is possibly the effect of the diffusion barrier formed by the fast crosslinking at the contact area of the alginate droplet and the injected fluid so that it obstructed the diffusion of more calcium ions into the drop. Similar heterogeneity has been observed in other external gelation approaches too [161,162,166,167], and integrating sodium chloride into the cross-linker calcium chloride was reported to be able to effectively improve the homogeneity [169], possibly the result of the competition between the two ions. Nevertheless, the cells were successfully trapped in the 3D micro scaffold.

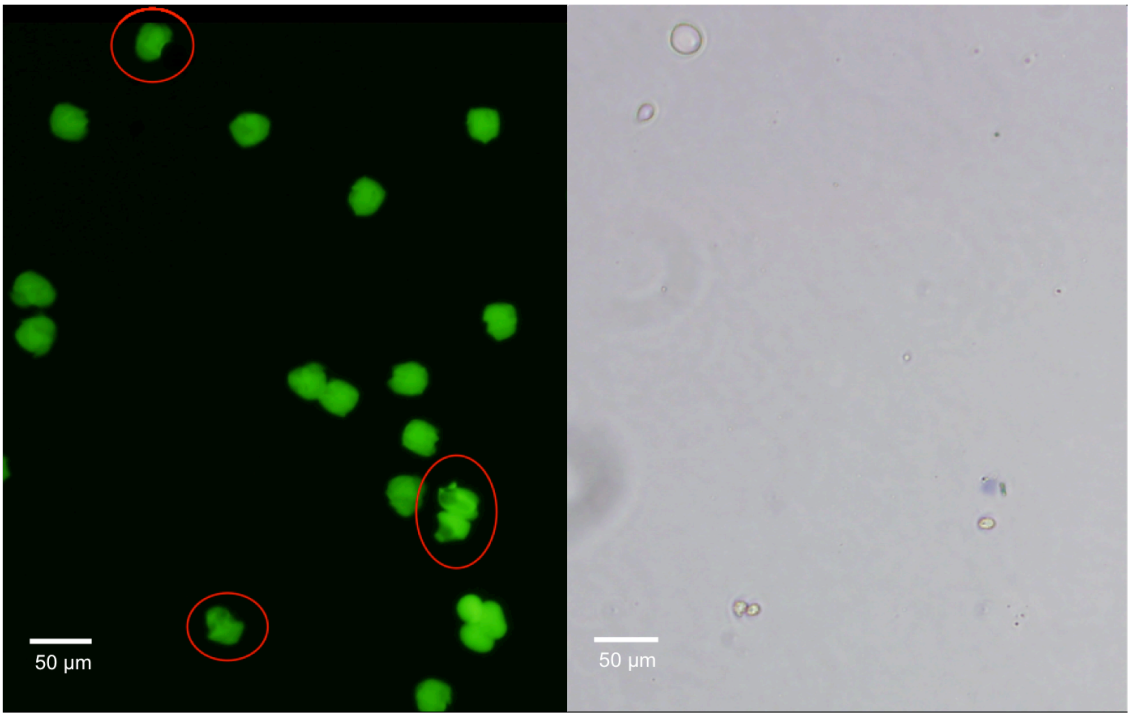
5.2.4 Viability analysis on the mMSCs cultured in the gel

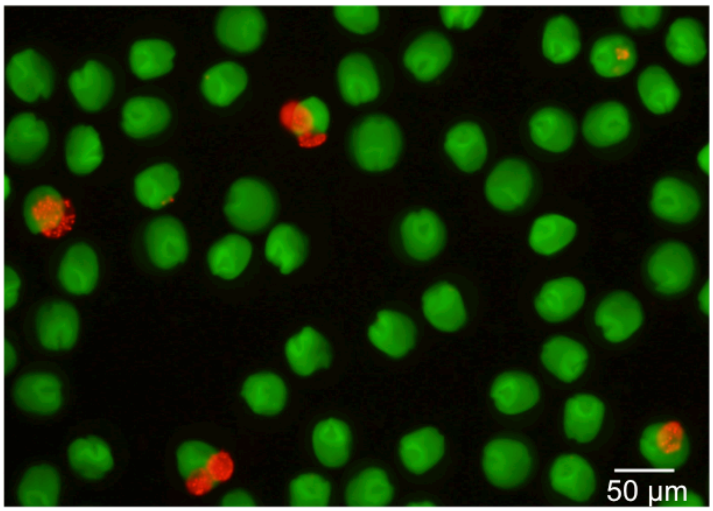
We cultured these gel beads in the cell medium to test the biocompatibility of the ECM synthesized with our microfluidic pico-injector-based method. On Day 3, a fraction of the beads were stained with Trypan blue and examined under confocal microscope and phase contrast microscope for cell viability characterization. We got a nearly 86% of viability in all the trapped cells examined.[3] Some cells even exhibited a suspected proliferation in the microscopic image (Figure 5.3b), where a group of cells were observed physically associated with each other in the immediate vicinity of the gel bead. As a comparison, a parallel experiment was performed on the gel beads synthesized from the internal gelation method where alginate polymers were encapsulated with calcium carbonate nanoparticles into the aqueous droplets surrounded by the 0.3% (v/v) acetic acid-incorporated oil phase. The viability assay showed no live cells present in the

[3] These are the cells that were not stained with PKH26; The viability was 25% for cells stained with PKH26.

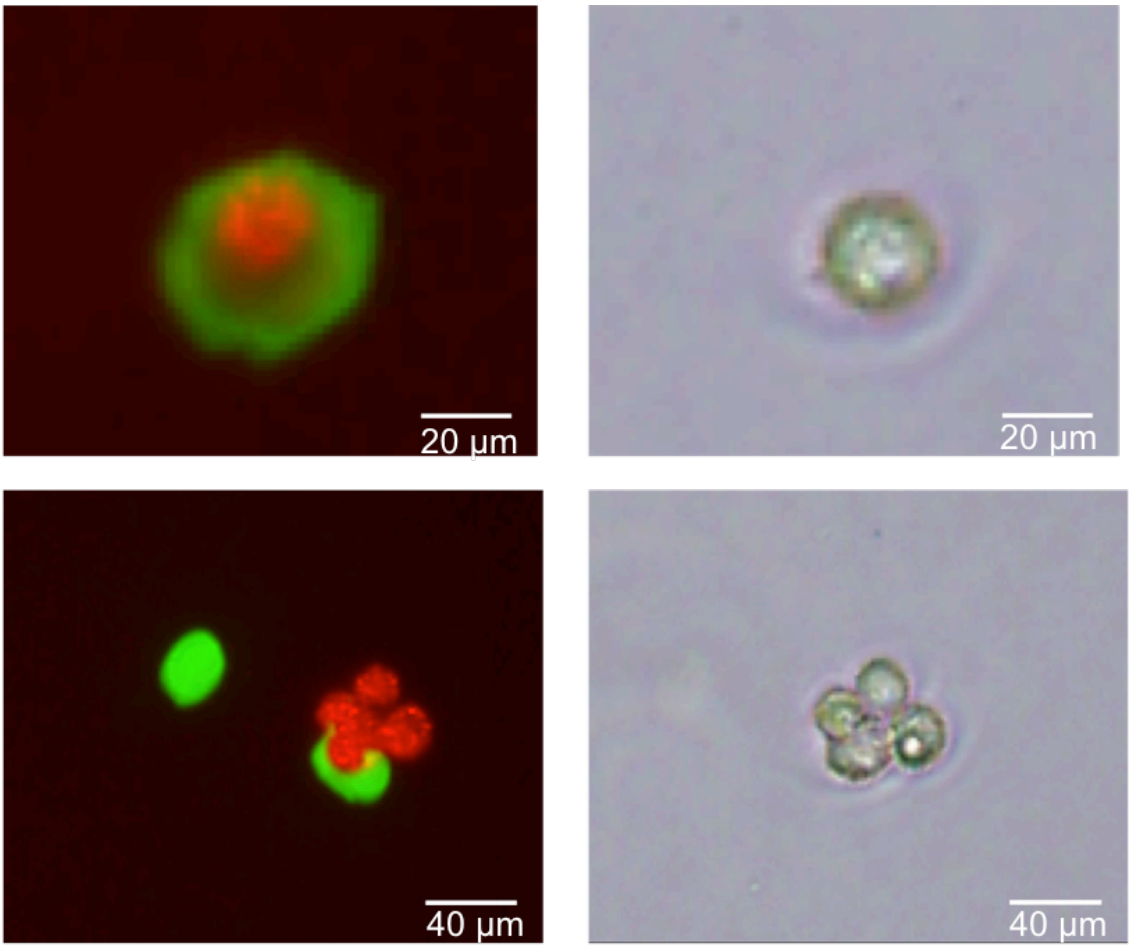
sampling gel beads on Day 3 (Figure 5.3c). Although more data on the longer-term culture are needed to elucidate the biocompatibility of the pico-injector approach by taking account of its effect on the cell proliferation and differentiation, the results we presented here at least suggested that our pico-injector approach is promising as a less-invasive 3D ECM synthesis method compared to the mild internal gelation means.

a. Day 0





b. Day 3



c. Parallel results from the internal gelation method (Day 3).

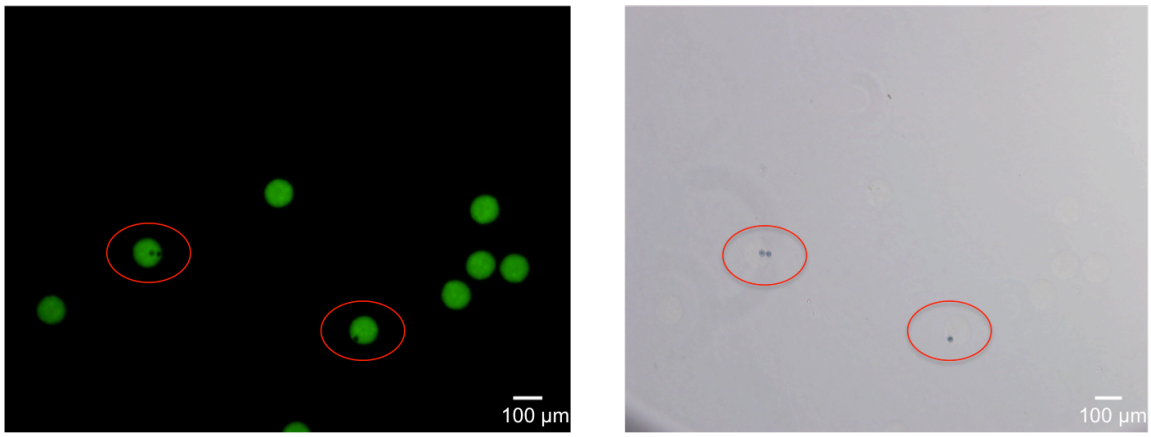


Figure 5.3 Cell viability measurement on the encapsulated mouse Mesenchymal stem cells. (a) Representative fluorescence (top left) and phase contrast (top right) micrographs of the cell-laden gel beads right after the gelation process. Red circles highlight the gel beads with the cells seeding on them. Trypan blue-based exclusion assay showed no dead cells in the examined field. The bottom picture is the superimposed fluorescence micrographs from the red and green channels, to show a clear view of the cells' location in the gel. (b) Fluorescence (left; superimposed from red and green channels) and phase contrast (right) micrographs of the live cells based on Trypan blue assay after 3 days' culture in the gel beads. Cells were stained with the red membrane dye PKH26 for easy identification. (c) Fluorescence (left) and phase contrast (right) micrographs of the cell-laden gel beads produced from the internal gelation method. Trypan blue assay was performed on Day 3. Red circles highlight the gel beads with the cells seeding on them.

5.3 Conclusion

We have successfully synthesized alginate gel beads on a pico-injector-based microfluidic platform that can reliably produce uniform gel beads down to 30 μm in size, nearly the lower limit as ECM for mammalian cells. By regulating the input cell density in the drop-maker and the flow rates in the pico-injector, we can have a better control on the encapsulated cell number and the gel composition. The distinctive high-throughput processing ability of the microfluidic platform ensures the efficiency in gel bead production. More importantly, the short-term cell viability assay suggests the mild invasiveness of this acid-free gelation approach on the D1 mouse Mesenchymal stem cells that seeded inside the synthesized 3D scaffolds, which is especially critical for stem cell niche studies that usually require the integrity of the cells for directed proliferation and differentiation under the regulation of the specific signaling.

For the long-term cell culture in gel beads, the targeted delivery of the specific nutrients and biochemical cues can be realized by employing a secondary pico-injection.

Moreover, the reiterative pico-injecting process can also be utilized to produce multi-core single cell-laden gel beads encapsulated within one common aqueous droplet and therefore sharing closer communications less interruptedly, which could hopefully facilitate the study of, for example, cell-cell interaction in a relatively isolated microenvironment.

Chapter 6

Conclusions

All the three projects described in this thesis share the same goal, which is to employ the highly sophisticated microfluidic platform to address biology-related problems, either to probe the fascinating fundamental questions (Chapter 5), or to seek the useful practical applications (Chapter 3, 4).

For the work in Chapter 3 and 4, where microfluidic drops were employed as the mobile carrier test tube, it exhibited the outstanding encapsulation ability in 'locking' the target and its metabolic product in a picon-size confinement. The target can be an active bacterial cell, or a mammalian cell with intricate metabolic activities going inside. Since cellular reactions are extremely crucial in determining macro-scale cell behaviors, studying these activities in the droplet could possibly enhance our understanding on the mechanisms underlying various fascinating or confusing cell behaviors.

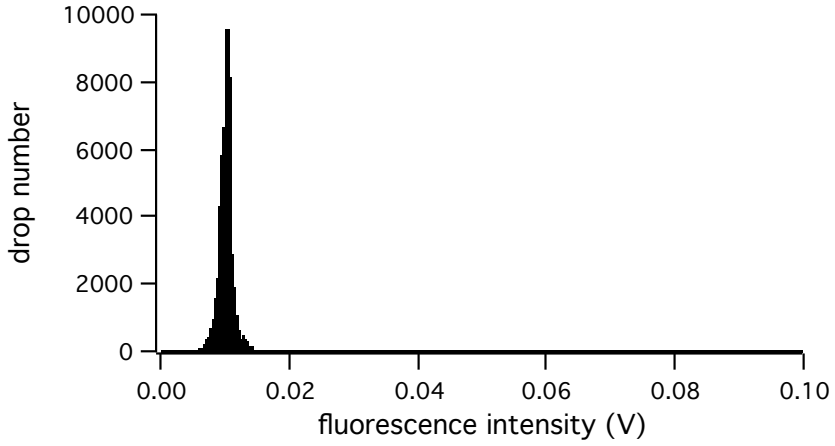
Although the drop dimension, ~20-50 μm in diameter, seems more matching the cellular level studies, the single DNA detection work illustrated in Chapter 3 indicates that the droplet, by incorporating the efficient in vitro transcription/translation (IVTT) reactions, can also serve as an ultrafast and highly sensitive tool for studies at the molecular level. As the essential building blocks for a series of complex metabolic processes, enzymatic reactions at the molecular level are always the aim of the sensitive detection technologies. Accurate and hopefully real-time detecting on the products or byproducts of the dynamic metabolic activities will guide us to a thorough understanding

of the cell behavior, and the potential in manipulating the participating biomolecules will confer us the capability in directing those enzymatic activities to function in our favorable way, for example, towards some particular clinical goal, both of which seem not impossible to be realized on the microfluidic platform with its continuing advancement.

Appendix

Supplementary Figures

a. λ = 0



b. λ = 10

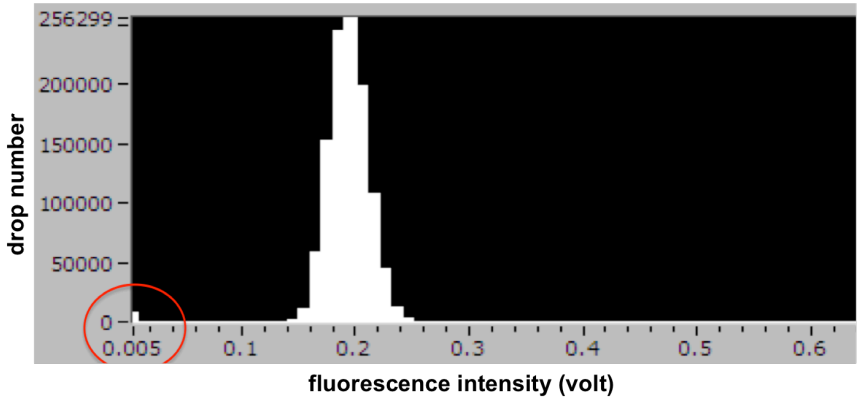
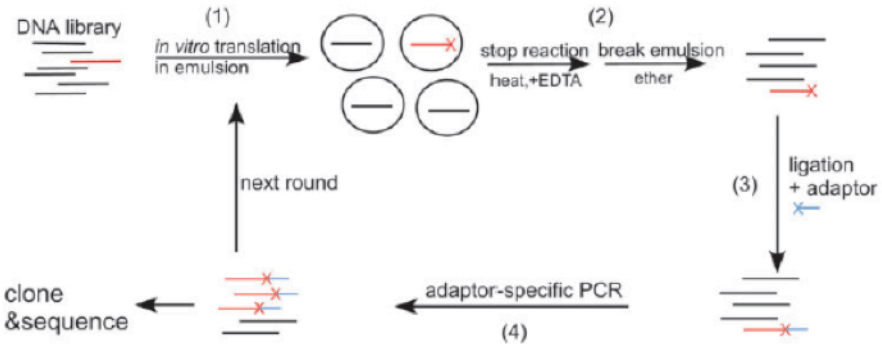


Figure S.1 Drop fluorescence intensity histogram from *cro-ad* detecting experiments: (a) λ = 0, (b) λ = 10. The red circle in (b) highlights the population with background fluorescence.

a.



b.

Calculation of theoretical enrichment

A Poisson distribution is assumed to describe the distribution of DNA templates in aqueous droplets. This implies that all droplets are of equal volumes. The probability that a droplet has $n = 0,1,2$ or more DNA templates can be calculated by:

$$f(n,\lambda) = \frac{e^{-\lambda}\lambda^n}{n!}$$

where λ is the ratio between the number of DNA templates and the number of droplets in the emulsion.

We assume that all the DNA templates in the same droplet containing at least one RE gene are selected, i.e. 100% selection efficiency. Thus the number of RE genes after selection is:

$$N_{re} = N\sum_{n=1} f(n,\lambda)\cdot\sum_{k=1}^{n}\binom{n}{k}\cdot p^k\cdot(1-p)^{n-k}\cdot k$$

where N is the total number of droplets in the emulsion and p is the percentage abundance of the RE gene in the starting library. The number of 'carryover' genes, which refer to those non-RE templates residing in the same droplets with an endonuclease gene, is:

$$N_{co} = N\sum_{n=1} f(n,\lambda)\cdot\sum_{k=1}^{n}\binom{n}{k}\cdot p^k\cdot(1-p)^{n-k}\cdot(n-k)$$

The final ratio between the selected RE gene and the 'carryover' gene, if we assume no bias in the PCR amplification, is:

$$r = \frac{N_{re}}{N_{co}}$$

The theoretical enrichment is approximately:

$$E = \frac{r}{p/(1-p)}$$

c.

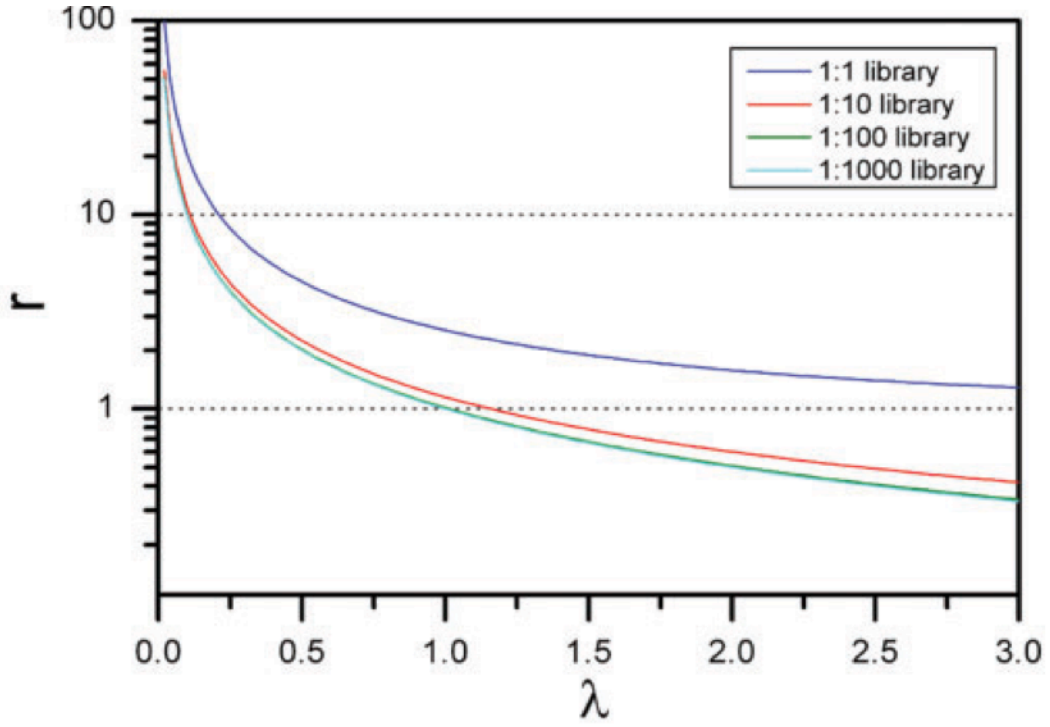


Figure S.2 Theoretical calculation for single-round sorting enrichment from Y. Zheng [107]. (a) Schematic of Yu's drop experiment where the calculation was based on. (b) Calculations given in Yu's paper. (c) Enrichment prediction curves for different λ values generated from the calculations in (b).

Bibliography

[1] J. C. McDonald, D. C. Duffy, J. R. Anderson, D. T. Chiu, H. Wu, O. J. Schueller, and G. M. Whitesides, Electrophoresis 21, 27 (2000).

[2] P. Garstecki, H. A. Stone, and G. M. Whitesides, Phys Rev Lett 94, 164501 (2005).

[3] A. S. Utada, E. Lorenceau, D. R. Link, P. D. Kaplan, H. A. Stone, and D. A. Weitz, Science 308, 537 (2005).

[4] T. Ward, M. Faivre, M. Abkarian, and H. A. Stone, Electrophoresis 26, 3716 (2005).

[5] S. Xu et al., Angew Chem Int Ed Engl 44, 724 (2005).

[6] D. R. Link, S. L. Anna, D. A. Weitz, and H. A. Stone, Phys Rev Lett 92, 054503 (2004).

[7] S. Koster et al., Lab Chip 8, 1110 (2008).

[8] J. Thiele, A. R. Abate, H. C. Shum, S. Bachtler, S. Forster, and D. A. Weitz, Small 6, 1723 (2010).

[9] A. R. Abate, J. Thiele, M. Weinhart, and D. A. Weitz, Lab Chip 10, 1774 (2010).

[10] A. R. Abate, J. Thiele, and D. A. Weitz, Lab Chip 11, 253 (2011).

[11] M. B. Romanowsky, M. Heymann, A. R. Abate, A. T. Krummel, S. Fraden, and D. A. Weitz, Lab Chip 10, 1521 (2010).

[12] A. R. Abate and D. A. Weitz, Small 5, 2030 (2009).

[13] A. R. Abate, T. Hung, P. Mary, J. J. Agresti, and D. A. Weitz, Proc Natl Acad Sci U S A 107, 19163 (2010).

[14] J. J. Agresti et al., Proc Natl Acad Sci U S A 107, 4004 (2010).

[15] J. C. Baret et al., Lab Chip 9, 1850 (2009).

[16] A. R. Abate, T. Hung, R. A. Sperling, P. Mary, A. Rotem, J. J. Agresti, M. A. Weiner, and D. A. Weitz, Lab Chip 13, 4864 (2013).

[17] Y. Zhou, H. Asahara, N. Schneider, P. Dranchak, J. Inglese, and S. Chong, J Am Chem Soc 136, 14031 (2014).

[18] J. A. Rowley, G. Madlambayan, and D. J. Mooney, Biomaterials 20, 45 (1999).

[19] J. W. Efcavitch and J. F. Thompson, Annu Rev Anal Chem (Palo Alto Calif) 3, 109 (2010).

[20] L. Ying, J. J. Green, H. Li, D. Klenerman, and S. Balasubramanian, Proc Natl Acad Sci U S A 100, 14629 (2003).

[21] M. Karymov, D. Daniel, O. F. Sankey, and Y. L. Lyubchenko, Proc Natl Acad Sci U S A 102, 8186 (2005).

[22] S. Hohng, R. Zhou, M. K. Nahas, J. Yu, K. Schulten, D. M. Lilley, and T. Ha, Science 318, 279 (2007).

[23] S. Dixit, M. Singh-Zocchi, J. Hanne, and G. Zocchi, Phys Rev Lett 94, 118101 (2005).

[24] R. T. Dame, M. C. Noom, and G. J. Wuite, Nature 444, 387 (2006).

[25] B. M. Reinhard, S. Sheikholeslami, A. Mastroianni, A. P. Alivisatos, and J. Liphardt, Proc Natl Acad Sci U S A 104, 2667 (2007).

[26] G. J. Wuite, S. B. Smith, M. Young, D. Keller, and C. Bustamante, Nature 404, 103 (2000).

[27] D. S. Johnson, L. Bai, B. Y. Smith, S. S. Patel, and M. D. Wang, Cell 129, 1299 (2007).

[28] G. Komazin-Meredith, R. Mirchev, D. E. Golan, A. M. van Oijen, and D. M. Coen, Proc Natl Acad Sci U S A 105, 10721 (2008).

[29] N. Hurt, H. Wang, M. Akeson, and K. R. Lieberman, J Am Chem Soc 131, 3772 (2009).

[30] N. A. Tanner, S. M. Hamdan, S. Jergic, K. V. Loscha, P. M. Schaeffer, N. E. Dixon, and A. M. van Oijen, Nat Struct Mol Biol 15, 170 (2008).

[31] N. A. Tanner, J. J. Loparo, S. M. Hamdan, S. Jergic, N. E. Dixon, and A. M. van Oijen, Nucleic Acids Res 37, e27 (2009).

[32] S. M. Hamdan, J. J. Loparo, M. Takahashi, C. C. Richardson, and A. M. van Oijen, Nature 457, 336 (2009).

[33] Y. Sako, S. Minoghchi, and T. Yanagida, Nat Cell Biol 2, 168 (2000).

[34] J. K. Jaiswal and S. M. Simon, Nat Chem Biol 3, 92 (2007).

[35] P. V. Cornish and T. Ha, ACS Chem Biol 2, 53 (2007).

[36] R. Roy, S. Hohng, and T. Ha, Nat Methods 5, 507 (2008).

[37] K. Gunther, M. Mertig, and R. Seidel, Nucleic Acids Res 38, 6526 (2010).

[38] M. Reuter and D. T. Dryden, Biochem Biophys Res Commun 403, 225 (2010).

[39] J. Zlatanova, S. M. Lindsay, and S. H. Leuba, Prog Biophys Mol Biol 74, 37 (2000).

[40] K. C. Neuman and A. Nagy, Nat Methods 5, 491 (2008).

[41] D. Pastre, L. Hamon, I. Sorel, E. Le Cam, P. A. Curmi, and O. Pietrement, Langmuir 26, 2618 (2010).

[42] A. Meller, L. Nivon, E. Brandin, J. Golovchenko, and D. Branton, Proc Natl Acad Sci U S A 97, 1079 (2000).

[43] A. Meller, L. Nivon, and D. Branton, Phys Rev Lett 86, 3435 (2001).

[44] S. Sorgenfrei, C. Y. Chiu, R. L. Gonzalez, Jr., Y. J. Yu, P. Kim, C. Nuckolls, and K. L. Shepard, Nat Nanotechnol 6, 126 (2011).

[45] T. R. Gibb, A. P. Ivanov, J. B. Edel, and T. Albrecht, Anal Chem 86, 1864 (2014).

[46] X. Gong, A. V. Patil, A. P. Ivanov, Q. Kong, T. Gibb, F. Dogan, A. J. deMello, and J. B. Edel, Anal Chem 86, 835 (2014).

[47] A. J. de Mello, Lab Chip 3, 29N (2003).

[48] M. Srisa-Art, A. J. deMello, and J. B. Edel, Chem Commun (Camb), 6548 (2009).

[49] S. Fields and O. Song, Nature 340, 245 (1989).

[50] S. L. Dove, J. K. Joung, and A. Hochschild, Nature 386, 627 (1997).

[51] J. K. Joung, E. I. Ramm, and C. O. Pabo, Proc Natl Acad Sci U S A 97, 7382 (2000).

[52] J. A. Hurt, S. A. Thibodeau, A. S. Hirsh, C. O. Pabo, and J. K. Joung, Proc Natl Acad Sci U S A 100, 12271 (2003).

[53] S. L. Dove and A. Hochschild, Cold Spring Harb Symp Quant Biol 63, 173 (1998).

[54] F. Courtois, L. F. Olguin, G. Whyte, D. Bratton, W. T. Huck, C. Abell, and F. Hollfelder, Chembiochem 9, 439 (2008).

[55] L. Mazutis et al., Anal Chem 81, 4813 (2009).

[56] L. Mazutis, J. C. Baret, P. Treacy, Y. Skhiri, A. F. Araghi, M. Ryckelynck, V. Taly, and A. D. Griffiths, Lab Chip 9, 2902 (2009).

[57] Y. Takeda, A. Sarai, and V. M. Rivera, Proc Natl Acad Sci U S A 86, 439 (1989).

[58] A. Fallah-Araghi, J. C. Baret, M. Ryckelynck, and A. D. Griffiths, Lab Chip 12, 882 (2012).

[59] C. O. Pabo, E. Peisach, and R. A. Grant, Annu Rev Biochem 70, 313 (2001).

[60] H. Wu, W. P. Yang, and C. F. Barbas, 3rd, Proc Natl Acad Sci U S A 92, 344 (1995).

[61] J. S. Kim and C. O. Pabo, Proc Natl Acad Sci U S A 95, 2812 (1998).

[62] P. Mary, V. Studer, and P. Tabeling, Anal Chem 80, 2680 (2008).

[63] F. Courtois, L. F. Olguin, G. Whyte, A. B. Theberge, W. T. Huck, F. Hollfelder, and C. Abell, Anal Chem 81, 3008 (2009).

[64] H. O. Smith and K. W. Wilcox, J Mol Biol 51, 379 (1970).

[65] K. Danna and D. Nathans, Proc Natl Acad Sci U S A 68, 2913 (1971).

[66] P. H. Johnson, A. S. Lee, and R. L. Sinsheimer, J Virol 11, 596 (1973).

[67] R. Zhang, Z. Zhu, H. Zhu, T. Nguyen, F. Yao, K. Xia, D. Liang, and C. Liu, Nucleic Acids Res 33, W489 (2005).

[68] J. N. Wolff and N. J. Gemmell, Biotechniques 44, 193 (2008).

[69] D. B. Stetson, J. S. Ko, T. Heidmann, and R. Medzhitov, Cell 134, 587 (2008).

[70] M. Wayengera, H. Kajumbula, and W. Byarugaba, Theor Biol Med Model 5, 18 (2008).

[71] J. T. Schiffer, M. Aubert, N. D. Weber, E. Mintzer, D. Stone, and K. R. Jerome, J Virol 86, 8920 (2012).

[72] N. Manjunath, G. Yi, Y. Dang, and P. Shankar, Viruses 5, 2748 (2013).

[73] P. Tebas et al., N Engl J Med 370, 901 (2014).

[74] R. J. Roberts, T. Vincze, J. Posfai, and D. Macelis, Nucleic Acids Res 43, D298 (2015).

[75] J. Eid et al., Science 323, 133 (2009).

[76] B. A. Flusberg, D. R. Webster, J. H. Lee, K. J. Travers, E. C. Olivares, T. A. Clark, J. Korlach, and S. W. Turner, Nat Methods 7, 461 (2010).

[77] Y. G. Kim, J. Cha, and S. Chandrasegaran, Proc Natl Acad Sci U S A 93, 1156 (1996).

[78] S. C. Ekker, Zebrafish 5, 121 (2008).

[79] J. A. Townsend, D. A. Wright, R. J. Winfrey, F. Fu, M. L. Maeder, J. K. Joung, and D. F. Voytas, Nature 459, 442 (2009).

[80] V. K. Shukla et al., Nature 459, 437 (2009).

[81] A. M. Geurts et al., Science 325, 433 (2009).

[82] A. Tovkach, V. Zeevi, and T. Tzfira, J Biotechnol 151, 1 (2011).

[83] J. F. Petolino, In Vitro Cell Dev Biol Plant 51, 1 (2015).

[84] M. Christian, T. Cermak, E. L. Doyle, C. Schmidt, F. Zhang, A. Hummel, A. J. Bogdanove, and D. F. Voytas, Genetics 186, 757 (2010).

[85] T. Li, S. Huang, W. Z. Jiang, D. Wright, M. H. Spalding, D. P. Weeks, and B. Yang, Nucleic Acids Res 39, 359 (2011).

[86] M. M. Mahfouz, L. Li, M. Shamimuzzaman, A. Wibowo, X. Fang, and J. K. Zhu, Proc Natl Acad Sci U S A 108, 2623 (2011).

[87] J. D. Sander, L. Cade, C. Khayter, D. Reyon, R. T. Peterson, J. K. Joung, and J. R. Yeh, Nat Biotechnol 29, 697 (2011).

[88] L. Tesson et al., Nat Biotechnol 29, 695 (2011).

[89] A. J. Wood et al., Science 333, 307 (2011).

[90] L. Cong, R. Zhou, Y. C. Kuo, M. Cunniff, and F. Zhang, Nat Commun 3, 968 (2012).

[91] N. E. Sanjana, L. Cong, Y. Zhou, M. M. Cunniff, G. Feng, and F. Zhang, Nat Protoc 7, 171 (2012).

[92] J. Guo, T. Gaj, and C. F. Barbas, 3rd, J Mol Biol 400, 96 (2010).

[93] V. Pattanayak, C. L. Ramirez, J. K. Joung, and D. R. Liu, Nat Methods 8, 765 (2011).

[94] R. Sorek, V. Kunin, and P. Hugenholtz, Nat Rev Microbiol 6, 181 (2008).

[95] H. Deveau, J. E. Garneau, and S. Moineau, Annu Rev Microbiol 64, 475 (2010).

[96] J. E. Garneau et al., Nature 468, 67 (2010).

[97] P. Horvath and R. Barrangou, Science 327, 167 (2010).

[98] G. Gasiunas, R. Barrangou, P. Horvath, and V. Siksnys, Proc Natl Acad Sci U S A 109, E2579 (2012).

[99] M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna, and E. Charpentier, Science 337, 816 (2012).

[100] L. Cong et al., Science 339, 819 (2013).

[101] P. Mali, L. Yang, K. M. Esvelt, J. Aach, M. Guell, J. E. DiCarlo, J. E. Norville, and G. M. Church, Science 339, 823 (2013).

[102] E. Pennisi, Science 341, 833 (2013).

[103] H. Wang, H. Yang, C. S. Shivalila, M. M. Dawlaty, A. W. Cheng, F. Zhang, and R. Jaenisch, Cell 153, 910 (2013).

[104] J. D. Sander and J. K. Joung, Nat Biotechnol 32, 347 (2014).

[105] R. M. Terns and M. P. Terns, Trends Genet 30, 111 (2014).

[106] E. R. Westra, A. Buckling, and P. C. Fineran, Nat Rev Microbiol 12, 317 (2014).

[107] Y. Zheng and R. J. Roberts, Nucleic Acids Res 35, e83 (2007).

[108] B. L. Wang, A. Ghaderi, H. Zhou, J. Agresti, D. A. Weitz, G. R. Fink, and G. Stephanopoulos, Nat Biotechnol 32, 473 (2014).

[109] K. R. Mesa et al., Nature (2015).

[110] S. Hasan, P. Hetie, and E. L. Matunis, Dev Biol (2015).

[111] C. K. Chan et al., Cell 160, 285 (2015).

[112] F. Gattazzo, A. Urciuolo, and P. Bonaldo, Biochim Biophys Acta 1840, 2506 (2014).

[113] L. Li and T. Xie, Annu Rev Cell Dev Biol 21, 605 (2005).

[114] A. Spradling, D. Drummond-Barbosa, and T. Kai, Nature 414, 98 (2001).

[115] B. R. Freedman, N. D. Bade, C. N. Riggin, S. Zhang, P. Haines, K. L. Ong, and P. A. Janmey, Biochim Biophys Acta (2015).

[116] B. Dong and S. Hayashi, Curr Opin Genet Dev 32, 129 (2015).

[117] R. Cai, T. Nakamoto, N. Kawazoe, and G. Chen, Biomaterials 52, 199 (2015).

[118] V. Akbarinejad, P. Tajik, M. Movahedin, R. Youssefi, S. Shafiei, and Z. Mazaheri, Anim Reprod Sci 157, 95 (2015).

[119] B. M. Baker and C. S. Chen, J Cell Sci 125, 3015 (2012).

[120] O. J. Tamplin, E. M. Durand, L. A. Carr, S. J. Childs, E. J. Hagedorn, P. Li, A. D. Yzaguirre, N. A. Speck, and L. I. Zon, Cell 160, 241 (2015).

[121] A. E. Shyer, T. R. Huycke, C. Lee, L. Mahadevan, and C. J. Tabin, Cell 161, 569 (2015).

[122] M. Huch et al., Cell 160, 299 (2015).

[123] C. C. Chen et al., Cell 161, 277 (2015).

[124] B. W. Carey, L. W. Finley, J. R. Cross, C. D. Allis, and C. B. Thompson, Nature 518, 413 (2015).

[125] R. C. Adam et al., Nature (2015).

[126] S. Ravindran and A. George, Front Physiol 6, 118 (2015).

[127] R. P. Pirraco and R. L. Reis, Stem Cell Rev (2015).

[128] M. Floren and W. Tan, Biomaterials 59, 39 (2015).

[129] N. T. Feric and M. Radisic, Adv Drug Deliv Rev (2015).

[130] F. De Francesco, G. Ricci, F. D'Andrea, G. F. Nicoletti, and G. A. Ferraro, Tissue Eng Part B Rev (2015).

[131] N. Daviaud, E. Garbayo, L. Sindji, A. Martinez-Serrano, P. C. Schiller, and C. N. Montero-Menei, Stem Cells Transl Med (2015).

[132] L. Conradi et al., Stem Cells Transl Med (2015).

[133] L. Zheng, X. Hu, Y. Huang, G. Xu, J. Yang, and L. Li, Biomed Mater 10, 015016 (2015).

[134] B. X. Zhang, Z. L. Zhang, A. L. Lin, H. Wang, M. Pilia, J. L. Ong, D. D. Dean, X. D. Chen, and C. K. Yeh, Tissue Eng Part A 21, 1611 (2015).

[135] B. Xu, H. Xu, Y. Wu, X. Li, Y. Zhang, X. Ma, and Q. Yang, PLoS One 10, e0124774 (2015).

[136] Q. Xing, K. Yates, M. Tahtinen, E. Shearier, Z. Qian, and F. Zhao, Tissue Eng Part C Methods 21, 77 (2015).

[137] E. Taghiabadi, S. Nasri, S. Shafieyan, S. Jalili Firoozinezhad, and N. Aghdami, Cell J 16, 476 (2015).

[138] S. A. Hurd, N. M. Bhatti, A. M. Walker, B. M. Kasukonis, and J. C. Wolchok, Biomaterials 49, 9 (2015).

[139] G. Chen et al., Biomaterials 52, 56 (2015).

[140] S. Chawla, S. Chameettachal, and S. Ghosh, Mater Sci Eng C Mater Biol Appl 49, 588 (2015).

[141] F. Zhang, K. Su, Y. Fang, S. Sandhya, and D. A. Wang, J Tissue Eng Regen Med 9, 77 (2015).

[142] J. Venkatesan, I. Bhatnagar, P. Manivasagan, K. H. Kang, and S. K. Kim, Int J Biol Macromol 72, 269 (2015).

[143] M. Szekalska, A. Amelian, and K. Winnicka, Acta Pharm 65, 15 (2015).

[144] R. S. Stowers, S. C. Allen, and L. J. Suggs, Proc Natl Acad Sci U S A 112, 1953 (2015).

[145] I. Sandvig, K. Karstensen, A. M. Rokstad, F. L. Aachmann, K. Formo, A. Sandvig, G. Skjak-Braek, and B. L. Strand, J Biomed Mater Res A 103, 896 (2015).

[146] A. Rajaram, D. J. Schreyer, and D. X. Chen, J Biomater Sci Polym Ed 26, 433 (2015).

[147] B. E. Larsen, J. Bjornstad, E. O. Pettersen, H. H. Tonnesen, and J. E. Melvik, BMC Biotechnol 15, 29 (2015).

[148] M. Komatsu, S. Konagaya, E. Y. Egawa, and H. Iwata, Biochim Biophys Acta (2015).

[149] S. H. Hong, M. Shin, J. Lee, J. H. Ryu, S. Lee, J. W. Yang, W. D. Kim, and H. Lee, Adv Healthc Mater (2015).

[150] J. W. Lee, Y. J. Park, S. J. Lee, S. K. Lee, and K. Y. Lee, Biomaterials 31, 5545 (2010).

[151] X. Cai, Y. Lin, G. Ou, E. Luo, Y. Man, Q. Yuan, and P. Gong, Cell Biol Int 31, 776 (2007).

[152] M. E. Burgi-Saville et al., Thyroid 8, 1147 (1998).

[153] D. Poncelet, Ann N Y Acad Sci 944, 74 (2001).

[154] L. Chan, H. Lee, and P. Heng, Int J Pharm 242, 259 (2002).

[155] C. P. Reis, R. J. Neufeld, S. Vilela, A. J. Ribeiro, and F. Veiga, J Microencapsul 23, 245 (2006).

[156] C. M. Silva, A. J. Ribeiro, D. Ferreira, and F. Veiga, Eur J Pharm Sci 29, 148 (2006).

[157] K. S. Huang, T. H. Lai, and Y. C. Lin, Front Biosci 12, 3061 (2007).

[158] K. Garcia-Gutierrez, H. M. Poggi-Varaldo, F. Esparza-Garcia, J. Ibarra-Rendon, and J. Barrera-Cortes, Bioprocess Biosyst Eng 34, 701 (2011).

[159] C. A. Hoesli, K. Raghuram, R. L. Kiang, D. Mocinecova, X. Hu, J. D. Johnson, I. Lacik, T. J. Kieffer, and J. M. Piret, Biotechnol Bioeng 108, 424 (2011).

[160] L. Liu, F. Wu, X. J. Ju, R. Xie, W. Wang, C. H. Niu, and L. Y. Chu, J Colloid Interface Sci 404, 85 (2013).

[161] D. Quong, R. J. Neufeld, G. Skjak-Braek, and D. Poncelet, Biotechnol Bioeng 57, 438 (1998).

[162] X. D. Liu, W. Y. Yu, Y. Zhang, W. M. Xue, W. T. Yu, Y. Xiong, X. J. Ma, Y. Chen, and Q. Yuan, J Microencapsul 19, 775 (2002).

[163] M. K. Das and P. C. Senapati, Acta Pol Pharm 64, 253 (2007).

[164] Y. Hu, Q. Wang, J. Wang, J. Zhu, H. Wang, and Y. Yang, Biomicrofluidics 6, 26502 (2012).

[165] G. Kaklamani, D. Cheneler, L. M. Grover, M. J. Adams, and J. Bowen, J Mech Behav Biomed Mater 36, 135 (2014).

[166] J. M. Puguan, X. Yu, and H. Kim, J Colloid Interface Sci 432, 109 (2014).

[167] J. Li, J. He, Y. Huang, D. Li, and X. Chen, Carbohydr Polym 123, 208 (2015).

[168] A. Schmitt et al., PLoS One 10, e0118937 (2015).

[169] C. Sargus-Patino, 2013).