MICROFLUIDIC ENGINEERING FOR ULTRASENSITIVE MOLECULAR ANALYSIS OF CELLS

by

Zhenning Cao

Dissertation submitted to the faculty of Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Biomedical Engineering

Chang Lu, Chair
Liwu Li
Sunghwan (Sunny) Jung
Rafael Davalos
Masoud Agah

August 28, 2015

Blacksburg, Virginia

Keywords: chromatin immunoprecipitation (ChIP), next generation sequencing, microfluidics, droplet sorting, protein translocation, sonication
The main focus of this research was the development of microfluidic technology for ultrasensitive and fast molecular analysis of cells.

Chromatin immunoprecipitation (ChIP) assay followed by next generation sequencing serves as the primary technique to characterize the genomic locations associated with histone modifications. However, conventional ChIP-seq assay requires large numbers of cells. We demonstrate a novel microfluidics-based ChIP-seq assay which dramatically reduced the required cell number. Coupled with next generation sequencing, the assay permitted the analysis of histone modifications at the whole genome from as few as ~100 cells. Using the same device, we demonstrated that MeDIP-seq with tiny amount of DNA (<5ng) generated high quality genome-wide profiles of DNA methylation.

Off-chip sonication often leads to sample loss due to multiple tube transferring. In addition, conventional sonicators are not able to manipulate samples with small volume(<100μl). We developed a novel microfluidic sonicator, which is able to achieve on-chip DNA/chromatin shearing into ideal fragment size (100~600bp) for both chromatin
immunoprecipitation (ChIP) and methylated DNA immunoprecipitation (MeDIP). The integrated on-chip sonication followed by immunoprecipitation (IP) reaction can significantly reduce sample loss and contamination.

Simple and accessible detection methods that can rapidly screen a large cell population with single cell resolution have been seriously lacking. We demonstrate a simple protocol for detecting translocation of native proteins using a common flow cytometer which detects fluorescence intensity without imaging. Using our approach, we successfully detected the translocation of native NF-κB (an important transcription factor) at its native expression level and examine the temporal dynamics in the process.

Droplets with encapsulated beads and cells have been increasingly used for studying molecular and cellular biology. However, a mixed population of droplets with an uneven number or type of encapsulated particles is resulted and used for screening. We developed a fluorescence-activated microfluidic droplet sorter that integrated a simple deflection mechanism. By passing droplets through a narrow interrogation channel, the encapsulated particles were detected individually. The microcontroller conducted the computation to determine the number and type of encapsulated particles in each droplet and made the sorting decision. Our results showed high efficiency and accuracy for sorting and enrichment.
ACKNOWLEDGEMENTS

First and foremost I would like to express my sincere gratitude to my advisor and head of my committee, Prof. Chang Lu, for his guidance and support. He gave me the opportunity to work on the exciting and challenging research projects and complete this dissertation in his lab. His support, guidance and encouragement in various ways enable my experience throughout the five year Ph.D. study to be productive and wonderful. I learned from him not only the specific research skills but also the critical thinking and problem solving skills. He also gave me many valuable advices on my career development. I believe that the experience working with him will benefit me in my life.

I would also like to extent my appreciation to Prof. Liwu Li, Prof. Sunghwan (Sunny) Jung, Prof. Rafael Davalos, and Prof. Masoud Agah for serving in my committee and providing insightful suggestions for my research work and dissertation. The completion of this dissertation would not have been possible without their time, patience and generous help.

I also owe my deepest gratitude to previous and current members in Dr. Lu’s group: Dr. Ning Bao, Dr. Tao Geng, Dr. Yihong Zhan, Dr. Despina Nelie Loufakis, Yousef Awwad, Chen Sun, Fangyuan Chen, Sai Ma, Hamid Hassanisaber, Yan Zhu, Travis Murphy and Mimosa Sarma for their help and friendship.

Last but not least, I am heartily thankful to my parents. Their love and devotion, and especially patience, are invaluable in all my endeavors.
Table of Contents

LIST OF TABLES ................................................................................................... xi
LIST OF FIGURES ................................................................................................ xii
LIST OF ABBREVIATIONS .................................................................................. xvii

CHAPTER 1. OVERVIEW OF EPIGENOMICS ....................................................... 1
  1.1 Eukaryotic chromatin structure ............................................................................... 1
  1.2 Chromatin modification .......................................................................................... 2
  1.3 Epigenetics ............................................................................................................ 3
  1.4 ChIP-seq ................................................................................................................ 4
    1.4.1 Chromatin immunoprecipitation ....................................................................... 5
    1.4.2 Next generation sequencing ............................................................................ 7
    1.4.3 Genome mapping ............................................................................................ 9
  1.5 Microfluidics for epigenetic Analysis ....................................................................... 9
    1.5.1 Screening of linearized chromatin................................................................. 10
    1.5.2 Chromatin immunoprecipitation in micro/nanofluidics ......................... 11

CHAPTER 2. A MICROFLUIDIC DEVICE FOR CHROMATIN IMMUNOPRECIPITATION FOLLOWED BY NEXT GENERATION SEQUENCING (ChIP-Seq) USING 100 CELLS .............................................................................................................................. 14
2.1 Introduction .................................................................................................................. 14

2.2 Materials and Methods .............................................................................................. 16

2.2.1 Fabrication of the microfluidic ChIP device .......................................................... 16

2.2.2 Setup of the microfluidic device .......................................................................... 19

2.2.3 Preparation of sonicated chromatin ................................................................. 20

2.2.4 Preparation of immunoprecipitation (IP) beads ................................................. 22

2.2.5 MOWChIP ............................................................................................................ 22

2.2.6 Extraction of ChIP DNA and input DNA ............................................................ 27

2.2.7 Construction of sequencing libraries .................................................................... 27

2.2.8 Cell culture .......................................................................................................... 28

2.2.9 Mouse strain, embryo dissection and cell sorting by fluorescence-activated cell sorting (FACS) ........................................................................................................... 28

2.2.10 MOWChIP-qPCR data analysis .......................................................................... 29

2.2.11 MOWChIP-Seq reads mapping and normalization ........................................... 30

2.2.12 Peak calling of MOWChIP-Seq data .................................................................. 31

2.2.13 Construction of Receiver Operating Characteristic (ROC) curves ................. 31

2.2.14 Correlation analysis of MOWChIP-Seq data with other published ChIP-Seq data sets ......................................................................................................................... 32

2.2.15 Prediction of enhancers and super enhancers using epigenomic data ............ 32

2.2.16 Transcription factor motif enrichment analysis ............................................... 33
### Chapter 2: Gene Ontology (GO) Term Enrichment Analysis of Super Enhancer Targets

2.2.17 Gene ontology (GO) term enrichment analysis of super enhancer targets 33

2.2.18 Assumptions of statistical tests 34

### Results and Discussion

2.3 Results and Discussion 34

2.3.1 Design and Operation of MOWChIP 34

2.3.2 Optimization of MOWChIP 35

2.3.3 Quantification of MOWChIP-seq using cell line cells 37

2.3.4 MOWChIP-seq of HSPCs 44

2.3.5 MOWChIP-seq using small number of cells 52

### Chapter 3: UltraSensitive Microfluidic Technology for Profiling Genome-Wide DNA Methylation

3.1 Introduction 57

3.2 Materials and methods 62

3.2.1 Fabrication of the microfluidic MeDIP device 62

3.2.2 Setup of the microfluidic device 63

3.2.3 Preparation of DNA sample 64

3.2.4 Preparation of immunoprecipitation (IP) beads 64

3.2.5 Microfluidic MeDIP 65

3.2.6 Elution of MeDIP DNA 66

3.2.7 Construction of sequencing libraries 66

3.2.8 Cell culture 67
3.2.9 Analysis of sequencing data ................................................................. 67

3.3 Results and discussion ........................................................................... 68

3.4 Conclusions ............................................................................................ 71

CHAPTER 4 ULTRA-FAST AND SENSITIVE ChIP/MeDIP ASSAY USING SONICATION INTEGRATED MICROFLUIDIC .............................................................. 72

4.1 Introduction .............................................................................................. 72

4.2 Methods and materials .......................................................................... 75

4.2.1 Fabrication of microfluidic device ....................................................... 75

4.2.2 Biological sample preparation ............................................................ 76

4.2.3 Setup of the sonication integrated microfluidic ChIP/MeDIP (SIM-ChIP/MeDIP) device ........................................................................................................ 77

4.2.4 Preparation of immunoprecipitation (IP) beads .................................. 78

4.2.5 SIM-ChIP/MeDIP ............................................................................... 78

4.2.6 Extraction of ChIP/MeDIP DNA and input DNA ................................. 80

4.2.7 Sample fragment size analysis using gel electrophoresis .................... 80

4.2.8 Real-time PCR .................................................................................... 80

4.3 Results and discussion .......................................................................... 81

4.3.1 Design of the microfluidic sonication device ..................................... 81

4.3.2 On-chip sonication of crosslinked cells and genomic DNA .............. 85

4.3.3 Integrated on-chip sonication followed by ChIP and MeDIP assay ..... 89
CHAPTER 5 Detecting intracellular translocation of native proteins quantitatively at the single cell level ................................................................. 97

5.1 Introduction ........................................................................................................ 97

5.2 Materials and methods ..................................................................................... 100

5.2.1 Cell sample preparation ............................................................................. 100

5.2.2 Standard fluorescence immunostaining ................................................... 100

5.2.3 Selective release of cytosolic proteins followed by fluorescence immunostaining ...................................................................................... 101

5.2.3 Fluorescence microscopy ......................................................................... 102

5.2.4 Western blotting ....................................................................................... 102

5.2.5 Flow cytometry analysis .......................................................................... 103

5.3 Results and Discussion ................................................................................. 103

5.4 Conclusions .................................................................................................... 113

CHAPTER 6 DROPLET SORTING BASED ON THE NUMBER/TYPe OF ENCAPSULATED PARTICLES USING A SOLENOID VALVE ...................... 115

6.1 Introduction .................................................................................................... 115

6.2 Materials and methods ................................................................................ 117

6.2.1 Microfluidic chip fabrication ................................................................ 117

6.2.2 Fluorescent beads and cell samples ....................................................... 118

6.2.3 System setup and operation .................................................................. 118
6.3 Results and discussion ................................................................. 120

6.4 Conclusions ................................................................................. 143

REFERENCES: ..................................................................................... 143

APPENDICES ....................................................................................... 171

Appendix A: Procedure for Microfluidic Chip Fabrication (Soft lithography) ............ 171

Appendix B: Cell Culture ........................................................................ 173

Appendix C Procedure for Quantitative Detection of Nucleocytoplasmic Transport. 175

Appendix D: Spectrum of Fluorophores.................................................... 188

Appendix E: MCU code for droplet sorting based on number of cell encapsulated in one droplet ......................................................... 189

Appendix F: MCU code for droplet sorting based on number and type of cell encapsulated in one droplet ......................................................... 193
LIST OF TABLES

TABLE 2. 1 PRIMER SEQUENCES USED IN MOWCHIP-QPCR. 35
TABLE 2. 2 PUBLISHED CHIP-SEQ DATA SETS USED IN THIS STUDY. 39
TABLE 2. 3 SUMMARY STATISTICS OF MOWCHIP-SEQ DATA GENERATED IN THIS STUDY. 44
TABLE 6. 1 RESULTS OF SORTING OUT TWO-BEAD AND TWO-CELL DROPLETS FROM MIXED DROPLET POPULATIONS. 135
TABLE 1 OF APPENDIX C TROUBLESHOOTING 187
LIST OF FIGURES

FIG. 1. 1 CHROMATIN STRUCTURE AND HISTONE MODIFICATIONS 2
FIG. 1. 2 KEY STEPS INVOLVED DURING THE STANDARD CHROMATIN IMMUNOPRECIPITATION (CHIP) 6
FIG. 2. 1 OVERVIEW OF THE MOWCHIP-SEQ PROTOCOL AND ITS OPTIMIZATION. 18
FIG. 2. 2 MOWCHIP-SEQ GENERATES HIGH QUALITY DATA USING AS FEW AS 100 CELLS. 40
FIG. 2. 3 EPIGENOMICS-AIDED DISCOVERY OF NOVEL ENHANCERS AND SUPER ENHANCERS IN FETAL LIVER HSPCS. 50
FIG. 2.S 1 SETUP OF THE MICROFLUIDIC DEVICE WITH ANCILLARY CONTROL AND LIQUID DELIVERY EQUIPMENT. 24
FIG. 2.S 2 MICROSCOPIC IMAGES THAT RECORDED THE MOWCHIP PROCESS. THE SCALE BAR REPRESENTS 1 MM. 26
FIG. 2.S 3 THE YIELD OF CHIP DNA AS A FUNCTION OF ANTIBODY CONCENTRATIONS USED DURING COATING OF THE IP BEADS. 37
FIG. 2.S 4 GENOME-WIDE CORRELATIONS AMONG MOWCHIP-SEQ DATA GENERATED USING VARIOUS GM12878 CELL SAMPLE SIZES. 42
FIG. 2.S 5 PURIFICATION OF HEMATOPOIETIC STEM AND PROGENITOR CELLS (HSPCS) FROM MOUSE FETAL LIVER. 46
FIG. 2.S 6 GENOME-WIDE CORRELATIONS BETWEEN MOWCHIP-SEQ DATA ON FL HSPCS DATA AND PUBLISHED DATA. 47
FIG. 2.S 7 SCATTER PLOTS OF HISTONE MODIFICATION SIGNALS AND TRANSCRIPT EXPRESSION LEVELS. FL LSK MICROARRAY DATA WAS DOWNLOADED FROM THE GEO DATABASE (ACCESSION NO.: GSE21404). 48

FIG. 2.S 8 NORMALIZED H3K27AC MOWCHIP-SEQ SIGNALS AT KNOWN ENHANCERS IN FL HSPCS. DATA GENERATED USING DIFFERENT NUMBERS OF CELLS ARE SHOWN. 51

FIG. 2.S 9 DNA FRAGMENT SIZE DISTRIBUTION AFTER SONICATION AND LIBRARY PREPARATION WITH STARTING CELL NUMBERS OF 600 AND 100. 53

FIG. 2.S10 GENOME-WIDE CORRELATIONS BETWEEN MOWCHIP-SEQ DATA GENERATED USING TWO PROTOCOLS FOR CHROMATIN PREPARATION. CHROMATIN WAS PREPARED FROM GM12878 CELLS. 54

FIG. 2.S 11 MOWCHIP-SEQ DATA USING 100 AND 600 STARTING CELLS. CHROMATIN WAS PREPARED FROM GM12878 CELLS. 55

FIG. 3. 1 THE PROCEDURE INVOLVED IN MEDIP-SEQ FOR EXAMINING GENOME-WIDE DNA METHYLATION. 61

FIG. 3. 2 SCHEMATIC ILLUSTRATION FOR THE FIVE MAJOR STEPS OF MICROFLUIDIC MEDIP. 69

FIG. 3. 3 DNA FRAGMENT SIZE DISTRIBUTION AFTER SONICATION. 70

FIG. 3. 4 PRELIMINARY MICROFLUIDIC MEDIP-SEQ DATA. 71

FIG. 4. 1 SCHEMATIC OF THE INTEGRATED SYSTEM INCORPORATING BOTH ON-CHIP SONICATION AND IP. 84

FIG. 4. 2 CAVITATION INSIDE THE MICROFLUIDIC CHAMBER. 85

FIG. 4. 3 THE PERFORMANCE OF ON-CHIP SONICATION. 88
FIG. 4. 4 THE PROCEDURE INVOLVED IN AN INTEGRATED MICROFLUIDIC SONICATION AND IP.
90
FIG. 4. 5 THE CORRELATION BETWEEN SONICATION TEMPERATURE AND CHIP RESULT.
93
FIG. 4. 6 QPCR RESULTS OF CHIP AND MEDIP AFTER ON-CHIP SONICATION.
95

FIG. 5. 1 THE COMPARISON BETWEEN FLOW CYTOMETRIC SCREENING OF CELLS AFTER (A) STANDARD FLUORESCENCE IMMUNOSTAINING AND (B) SELECTIVE RELEASE OF CYTOSOLIC PROTEINS FOLLOWED BY IMMUNOSTAINING.
104

FIG. 5.2 THE DETECTION OF NF-KB TRANSLOCATION USING THE SELECTIVE RELEASE/IMMUNOSTAINING PROTOCOL AND CONVENTIONAL FLOW CYTOMETRY.
106

FIG. 5. 3 THE DOSE DEPENDENCE OF TNFα STIMULATION ANALYZED BY OUR APPROACH.
111

FIG. 5. 4 THE TEMPORAL DYNAMICS OF NF-KB TRANSLOCATION DETECTED BY OUR APPROACH (A) AND VERIFIED BY WESTERN BLOTTING ANALYSIS OF THE NUCLEAR FRACTION (B).
112

FIG. 5.S 1 FLUORESCENCE IMAGES OF UNTREATED CELLS (A, C) AND TNFα-STIMULATED CELLS (B, D) AFTER STANDARD IMMUNOSTAINING (A, B) AND OUR SELECTIVE-RELEASE-BASED IMMUNOSTAINING (C, D).
108

FIG. 5.S 2 THE OPTIMIZATION OF CONDITIONS (SAPONIN CONCENTRATION AND TREATMENT DURATION) FOR MAXIMAL DIFFERENTIATION OF UNTREATED
AND TNFα STIMULATED CELLS USING SELECTIVE-RELEASE-BASED IMMUNOSTAINING AND FLOW CYTOMETRY. 109

FIG. 6. 1 SCHEMATIC OF THE DROPLET SORTING SYSTEM. 123
FIG. 6. 2 DROPLET SORTING BASED ON EMITTED FLUORESCENCE. 126
FIG. 6. 3 THE FLUORESCENCE SIGNAL CONTAINS INFORMATION ON THE NUMBER OF ENCAPSULATED PARTICLES PER DROLET. 128
FIG. 6. 4 THE PROCESSING OF THE FLUORESCENCE INTENSITY DATA BY THE COMPARATOR. 130
FIG. 6. 5 A FLOWCHART THAT SUMMARIZES THE LOGIC OF THE MICROCONTROLLER FOR SORTING TWO-PARTICLE DROPLETS OUT OF A MIXED POPULATION. 132
FIG. 6. 6 CAMERA IMAGES SHOW THAT ONLY TWO-PARTICLE DROPLETS WERE SORTED OUT FROM A MIXED POPULATION. 134
FIG. 6. 7 SCHEMATIC OF DIFFERENTIAL LABELING STRATEGY. 137
FIG. 6. 8 SCHEMATIC OF THE MODIFIED DROPLET SORTING SYSTEM FOR SORTING OUT HETEROTYPIC CELL PAIR ENCAPSULATED DROPLET. 138
FIG. 6. 9 THE FLUORESCENCE SIGNAL CONTAINS INFORMATION ON THE TYPE AND NUMBER OF ENCAPSULATED CELLS PER DROPLET. 140
FIG. 6. 10 A FLOWCHART THAT SUMMARIZES THE LOGIC OF THE MICROCONTROLLER FOR SORTING HETEROTYPIC CELL-PAIR DROPLETS OUT OF A MIXED POPULATION. 141
FIG. 6. 11 CAMERA IMAGES CONFIRM THE SUCCESSFUL ISOLATION OF HETEROTYPIC CELL-PAIR FROM A MIXED DROPLET POPULATION. 142
FIG. 6S. 1 THE SIMULATION OF THE DEFLECTION PROCESS BY PRESSURE ACTUATION BY COMSOL MULTIPHYSICS 4.2. 122

FIG. 6S. 2 DIAGRAM OF THE COMPARATOR CIRCUIT. 129
# LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>SYMBOL</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>AC</td>
<td>Alternative current</td>
</tr>
<tr>
<td>BioMEMS</td>
<td>Biological micro-electromechanical system</td>
</tr>
<tr>
<td>CHO</td>
<td>Chinese hamster ovary</td>
</tr>
<tr>
<td>ChIP</td>
<td>Chromatin immunoprecipitation</td>
</tr>
<tr>
<td>DC</td>
<td>Direct current</td>
</tr>
<tr>
<td>De</td>
<td>Dean</td>
</tr>
<tr>
<td>DI</td>
<td>Dionized</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>dsDNA</td>
<td>Double-stranded Deoxyribonucleic acid</td>
</tr>
<tr>
<td>EDC</td>
<td>EDTA-carbodiimide</td>
</tr>
<tr>
<td>EFC</td>
<td>Electroporative flow cytometry</td>
</tr>
<tr>
<td>EGFP</td>
<td>Enhanced green fluorescent protein</td>
</tr>
<tr>
<td>FCS</td>
<td>Fluorescence correlation spectroscopy</td>
</tr>
<tr>
<td>FI</td>
<td>Fusion index</td>
</tr>
<tr>
<td>Acronym</td>
<td>Full Form</td>
</tr>
<tr>
<td>---------</td>
<td>-----------</td>
</tr>
<tr>
<td>FWHM</td>
<td>Full width at half maximum</td>
</tr>
<tr>
<td>GFP</td>
<td>Green fluorescent protein</td>
</tr>
<tr>
<td>IKK</td>
<td>IκB kinase</td>
</tr>
<tr>
<td>IL-1 β</td>
<td>Interleukin-1 β</td>
</tr>
<tr>
<td>LSC</td>
<td>Laser scanning cytometry</td>
</tr>
<tr>
<td>MNase</td>
<td>Micrococcal nuclease</td>
</tr>
<tr>
<td>NFκB</td>
<td>Nuclear factor-κB</td>
</tr>
<tr>
<td>NLS</td>
<td>Nuclear localization sequence</td>
</tr>
<tr>
<td>NPC</td>
<td>Nuclear pore complex</td>
</tr>
<tr>
<td>PDMS</td>
<td>Poly(dimethylsiloxane)</td>
</tr>
<tr>
<td>PMT</td>
<td>Photomultiplier</td>
</tr>
<tr>
<td>PI</td>
<td>Propidium iodide</td>
</tr>
<tr>
<td>Re</td>
<td>Reynolds number</td>
</tr>
<tr>
<td>Rc</td>
<td>Channel Reynolds number</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>Rp</td>
<td>Particle Reynolds number</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>------------------------------------------------</td>
</tr>
<tr>
<td>RSD</td>
<td>Relative standard deviation</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulphate</td>
</tr>
<tr>
<td>ssDNA</td>
<td>Single-stranded deoxyribonucleic acid</td>
</tr>
<tr>
<td>Syk</td>
<td>Spleen tyrosine kinase</td>
</tr>
<tr>
<td>TNFα</td>
<td>Tumor necrosis factor alpha</td>
</tr>
<tr>
<td>u</td>
<td>Kinematic viscosity</td>
</tr>
<tr>
<td>Vp</td>
<td>Processed signal</td>
</tr>
<tr>
<td>Vref</td>
<td>Reference signal</td>
</tr>
</tbody>
</table>
CHAPTER 1. OVERVIEW OF EPIGENOMICS

1.1 Eukaryotic chromatin structure

In eukaryotic cells, nearly two meters of DNA are packaged with histones into chromosomes. Chromosome are made up of chromatin, the complex of DNA and histones. A series of hierarchical structures are involved during the packaging and the lowest level of this kind of structure occurs through the wrapping of DNA into a nucleosome. A nucleosome consists of 146 base pairs (bps) of DNA wrapping a protein octamer. The protein octamer is composed of eight subunits called histones. There are five different types of histones. Four of them are called core histones, named H2A, H2B, H3, and H4. Each type of core histone has two copies to form a nucleosome core particle. The fifth type of histone is called H1. It functions to stabilize the nucleosome structure by binding at the linker DNA region, whose length can range from several kilobase pairs to as small as several base pairs [1]. A long DNA molecule is packaged into a series of nucleosomes, which are connected to one another by a short region of DNA, and form ‘beads on a string’ fiber [2]. The ‘beads-on-a-string’ model describes the primary structure of DNA packaging. At the next level of packaging, multiple histones wrap into a 30 nm fiber consisting of nucleosome arrays in their most compact form. The chromatin fiber is then further folded into higher-order structures (metaphase chromosome), which is still not fully understood and actively debated [3, 4].
1.2 Chromatin modification

The primary structure of a histone is a long chain of 20 types of amino acids, which can be referred as a residue. The underlying residues of four types of core histones can be subjected to a variety of modifications that alter DNA accessibility for transcription factors and other molecules such as RNA polymerase [5]. Such modification refers to a chemical attachment of a functional group to a histone residue, as well as their different forms (mono-, di-, trimethylation for example), by covalent bonds [6]. Because such reactions usually take place after the protein sequence is translated from mRNA, they are also called post-translational modifications. The variety of known histone modifications is growing [7]. The most extensively documented modifications include acetylation, methylation, phosphorylation, ubiquitination, and biotinylation, named after the different
functional groups attached [8]. We denote by H3K4ac a histone modification, i.e., lysine (K) acetylation (ac) at the 4th residue of histone H3. Some modifications can involve more than one functional group, e.g., H3K4 has monomethylation (attached with one methyl group), dimethylation (with two methyl groups), and trimethylation (with three methyl groups); we denote by H3K4me1, H3K4me2, and H3K4me3, respectively. In particular, acetylation on histone H4 and methylation of lysine 9 on histone H3 can significantly influence chromatin structure of euchromatin (decondensed chromatin) and heterochromatin (highly condensed chromatin), respectively[9-12]. Hence, patterns of these histone modifications, collectively referred to as the “histone code” can physically regulate gene expression and transcription [13]. The histones can also be exchanged with variants (such as H2A.X, H2A.Z, H3.3) and the modifications and variants have both been associated with many functional responses [14].

Generally, chromatin modifications also include DNA methylation. DNA methylation is defined as the addition of a methyl group to the cytosine in a CpG dinucleotide [5]. This covalent modification is known to have important functions in gene regulation and chromosomal stability and parental imprinting [5]. CpG dinucleotides are generally underrepresented in eukaryotic genomes but most (60–80%) are methylated [15].

1.3 Epigenetics

The genetic information of an organism is heritable through DNA replication, a process in which DNA is replicated during cell division [16]. The DNA sequence of a protein-coding gene uniquely determines the type of protein to which it is translated. The biological information flow from DNA sequence to RNA to protein in the processes of
replication, transcription, and translation is called the ‘central dogma’ of molecular biology [17]. Epigenetics, however, is the study of stable and heritable changes in gene expression without altering the underlying DNA sequence. In plainer language, epigenetics is the study of changes in the expression of genes caused by certain base pairs in DNA, or RNA, being "turned off" or "turned on" again, through chemical reactions. There are several intricately related forms of epigenetic mechanisms including aforementioned DNA methylation, histone modifications, histone variants, as well as chromatin remodeling, genomic imprinting, and RNA interference (RNAi) [18, 19].

These epigenetic mechanisms play an essential role in the development of normal cells as well as the regulation of biological processes such as aging and memory formation [20]. Alternation in epigenetic regulations could give rise to various cancers and syndromes [20, 21]. The increasing knowledge of epigenetics has enabled the applications of epigenetic therapy in clinical setting. Some inhibitors of chromatin modifying enzymes, particularly histone deacetylases and DNA methyltransferases, are currently being tested in clinical trial for the treatment of cancers [22-24].

1.4 ChIP-seq

The study of epigenetics of higher organisms including human and mouse is made possible by the recent advances in experimental technique ChIP-Seq, chromatin immunoprecipitation (ChIP) combined with high-throughput sequencing (seq). The basic idea of ChIP-Seq is to find the DNA fragments that are associated with the target epigenetic mark and to determine the sequences of the DNA fragments. Then, one can
identify the genomic locations of the epigenetic mark by comparing the DNA sequences obtained from the chromatin sample with those in a reference genome database.

The epigenetic mark detected using ChIP-Seq can be a particular histone modification, histone variant, or a DNA-binding protein (e.g. transcription factor). The experimental procedures in the ChIP step are different for different epigenetic marks. Here, we use histone modification ChIP-Seq as a representative to describe the ChIP-Seq technique. Briefly, the strategy of ChIP-Seq consists of three major parts: chromatin immunoprecipitation (ChIP), high-throughput sequencing (Seq), and genomic mapping.

1.4.1 Chromatin immunoprecipitation

Chromatin Immunoprecipitation (ChIP) is a type of immunoprecipitation experimental technique used to detect protein binding on a region of DNA. It can determine whether specific proteins (e.g. transcription factors) are biologically associated with specific genomic regions (e.g. enhancer, promoter). In addition, ChIP can determine the specific location of genome where different types of histone modification happen. Briefly, this method is as follows: intermolecular interactions, including protein-DNA interactions, are stabilized by formaldehyde cross-linking, and are then sonicated into 200~700bp chromatin fragments (cross-linked ChIP, XChIP). Alternatively, chromatin without crosslinking can be fragmented by micrococcal nuclease digestion (Native ChIP, NChIP). The previous fragmentation technique is suited for mapping the DNA target of transcription factors. While the later one is suited for mapping the DNA target of histone modifiers, as histones and its wrapped DNA are naturally linked. At this point, the immunoprecipitation, which is the technique of precipitating a target antigen out of
solution using a particular antibody to bind, is performed resulting in the concentration and purification of target protein-DNA complexes. The protein–DNA complexes are then heated to reverse the cross-linking of the protein and DNA complexes, allowing the DNA to be separated from the proteins. The quantity of the purified DNA fragments can then be identified by polymerase chain reaction (PCR), microarrays (ChIP-chip) and sequencing (ChIP-seq). The limitation of PCR assay is that one must have an idea which genomic region is being targeted in order to generate the correct PCR primers. Alternatively, DNA microarray (ChIP-chip) can be used allowing for the characterization of protein binding on a genome-wide scale. Recently, the application of next generation sequencing (NGS) to analyze the immunoprecipitated DNA has replaced ChIP-chip as the preferred means of harvest genome-wide data, and confers a number of advantages including higher resolution, improved quantification range, greater genome coverage, fewer artifacts, and lower cost [25].

**Fig. 1.** Key steps involved during the standard chromatin immunoprecipitation (ChIP)
1.4.2 Next generation sequencing

Nucleic acid sequencing is a method for determining the precise order of nucleotides within a given DNA or RNA molecule. The first major foray into DNA sequencing was the Human Genome Project, a $3 billion, 13-year-long endeavor, completed in 2003. The Human Genome Project was accomplished with first-generation sequencing, known as Sanger sequencing. Sanger sequencing (the chain-termination method), developed in 1975 by Edward Sanger, was considered the gold standard for nucleic acid sequencing for the subsequent two and a half decades[26]. Since completion of the first human genome sequence, a variety of new DNA-sequencing platforms have emerged, as a result of the increasingly demand for cheaper and faster sequencing methods. These new types of sequencing techniques are commonly referred to as next-generation sequencing (NGS) technologies. The primary difference between these platforms and the Sanger capillary sequencing is the drastic increased throughput achieved through the massively parallel sequencing, during which millions of fragments of DNA from a single sample are sequenced in unison. With its fast development, next-generation sequencing applies to whole genome sequencing, transcriptome profiling (RNA-Seq), DNA-protein interactions (ChIP-sequencing), and epigenome characterization [27].

Most commonly used NGS platforms are 454 pyrosequencing, Illumina’s cluster based sequencing, ABI SOLiD sequencing, ion torrent semiconductor sequencing and DNA nanoball sequencing. Among them, Illumina provide industry-leading sequencing technology in respects of data quality, throughput and cost. The principle of Illumina’s
sequencing technique is based on the sequencing-by-synthesis chemistry, with novel reversible terminator nucleotides for the four bases each labelled with a different fluorescent dye, and a special DNA polymerase enzyme able to incorporate them [28]. DNA fragments are ligated at both ends to adapters and, after denaturation, immobilized at one end on a solid support. The surface of the support is coated densely with the adapters and the complementary adapters. Each single-stranded fragment, immobilized at one end on the surface, creates a ‘bridge’ structure by hybridizing with its free end to the complementary adapter on the surface of the support. In the mixture containing the PCR amplification reagents, the adapters on the surface act as primers for the following PCR amplification. Again, amplification is needed to obtain sufficient light signal intensity for reliable detection of the added bases. After several PCR cycles, random clusters of about 1000 copies of single-stranded DNA fragments (termed DNA ‘polonies’, resembling cell colonies after polymerase amplification) are created on the surface. The reaction mixture for the sequencing reactions and DNA synthesis is supplied onto the surface and contains primers, four reversible terminator nucleotides each labelled with a different fluorescent dye and the DNA polymerase. After incorporation into the DNA strand, the terminator nucleotide, as well as its position on the support surface, is detected and identified via its fluorescent dye by the CCD camera. The terminator group at the 3'-end of the base and the fluorescent dye are then removed from the base and the synthesis cycle is repeated. The sequence read length achieved in the repetitive reactions is about 50 nucleotides. The sequence of at least 40 million polonies can be simultaneously determined in parallel, resulting in a very high sequence throughput, on the order of Gigabases per support.
1.4.3 Genome mapping

The final part is to find the genomic locations of the tags. This requires a known reference genome for the species under investigation. This step involves the alignment between the sequence tag and the reference genome sequence. The key issue in the alignment is whether the tag can be uniquely mapped to the reference genome. The alignment feasibility can be estimated by a probabilistic model. If the length of the tag is longer than 20 bps, it is expected that it can be uniquely mapped to the reference genome. Of course, the actual genome is not random and contains repetitive regions. As a result, there are tags that are mapped to multiple locations into the genome, which are usually discarded. In summary, millions of tags are mapped to the genome by such alignment. As a result, the DNA sequence of each tag is replaced with the genomic coordinates.

1.5 Microfluidics for epigenetic Analysis

Although current ChIP-related assays have been generating useful data, the technique has some serious limitations. First, a key limitation is the requirement of a large number of cells (>10^6 cells per IP for ChIP-qPCR and 10^7-10^8 cells for ChIP-seq). This is usually feasible with cell lines but poses a serious challenge when primary cells are used. The sample amount generated by lab animals and patients is very limited. For example, the number of naturally occurring T regulatory cells in murine splenocytes is ~ 10,000 per spleen, and ~5000 per ml peripheral blood leukocyte. Circulating tumor cells are present by the frequency of 1-10 per ml of whole blood in patients with metastatic cancer. In addition, primary samples typically contain a mixture of different cell types. The enrichment and isolation of a homogenous single cell type not only add time and labor to
the protocol but also generate further loss in the sample amount. To date, efforts to improve conventional ChIP have largely been successful in reducing sample to thousands of cells per assay (e.g. ~1000 cells for ChIP-qPCR, and ~10,000 cells, involving whole genome amplification, for ChIP-seq) [29-38], but have not provide any genome-wide profile of epigenetic modification from less than 10000 cells. Second, most ChIP assays involve extensive manual handling [39, 40] and take 3-4 days or longer to finish. These cumbersome procedures may create loss of materials and technical errors that lead to inconsistencies between replicates.

To address the two aforementioned major challenges in scaling up epigenetic screening, recent advancements in micro/nanotechnology offer opportunities for scalable, low consumption epigenetic screening techniques, using mechanisms that allow for precise control, improved handling, and higher throughput studies.

1.5.1 Screening of linearized chromatin

Since chromatin exists in a highly compact state, specific markers in chromatin are quite difficult to discriminate when tangled. A straightforward method is to directly observe the position of epigenetic markers along a strand of linearized chromatin. In order to visualize single chromatin molecule and detect target epigenetic marker on it, single chromatin fiber with its associated protein needs to be stretched intact. For this purpose, a variety of micro/nano-scale techniques, including molecular combing nanoconfinement, hydrodynamic squeezing, have been introduced and developed [41-44]. After the DNA or chromatin unfolded, direct imaging of fluorescently tagged based modifications on individual DNA/chromatin molecule can be done, via methods such as fluorescence in
situ hybridization (FISH). It was reported that, with fluorescent labeling using multiple antibodies against specific histone modification, epigenetic modified DNA molecular can be detected, sorted and recovered toward subsequent DNA sequencing \[42, 45, 46\]. In spite of its great potential, direct screening and characterization of linearized chromatin has several limitations for epigenetic study. First, this set of approaches provide limited information about exact sequence location of epigenetic modifications as only “fingerprint” patterns of chromatin/DNA fiber are produced. Second, the resolution of such techniques is in general limited by diffraction-limited resolution (~200nm or ~588bp), which is insufficient to resolve modifications that are in close proximity \[46, 47\]. Although this limitation may be ameliorated with super-resolution microscopy, it has limited accessibility to ordinary labs at this stage. Third, the ability to stretch DNA or even difficult chromatin gently and homogenously is still limited. As a result, chromatin or DNA cannot be fully oriented for researcher to obtain a precise map of the positions of DNA or chromatin modifications, consequently inhibiting the final resolution \[43\] \[44\]. These noted challenges have led to development of alternative micro/nano scale techniques to identify histone modifications.

1.5.2 Chromatin immunoprecipitation in micro/nanofluidics

As noted, conventional ChIP for small number of cells can be challenging to perform, mostly because the low efficiency during immunoprecipitation(IP). In addition, the sequential washing and purification steps after immunoprecipitation involve multiple and tedious manual steps, which may lead to significant user variability, increased assay time and excessive sample loss.
In contrast to conventional technologies, miniaturization of the IP process in a microfluidic format would offer a number of benefits. In particular, it would make IP much more efficient, as the close proximity among antigen and antibody, inside microfluidic chamber/channel (several pl~nl), can greatly reduce the diffusion length of small molecular and promote reaction rate of IP. In addition, microfluidics is amenable to applications where only small numbers of cells are available (such as biopsy samples, aged cells, rare primary cells and purified stem cells). Furthermore, microfluidic approach would also enable upstream sample preparation components, such as cell lysis devices [48], as well as downstream multiplex screening or analysis components, such as on-chip-qPCR or on-chip gel electrophoresis [49-52]. This kind of integration would further minimize hands-on time and sample losses.

There have been a number of reports describing immunoprecipitation assay on various formats of microfluidic platforms. Functionalized surface of micro-chambers or micro-channels is most commonly used to capture target protein molecules [53-58]. Recently, functionalized microbeads have been applied inside microfluidic channel to immobilized target biomolecules. By introducing micro-beads (usually around several micrometer in diameter), e.g. micro magnetic beads, the surface area to volume ration inside microfluidic channel can be greatly promoted, which enhances the available binding sites per unit volume. As a result, more target molecules can be immobilized in a limited time comparing with methods via functionalized channel surface. Therefore, the detection limit can be dramatically increased. In addition, the presence of beads (especially packed beads) in microchannel will further reduce the diffusion distance of analytes to their receptors, thus shortening the analysis time and increasing the sensitivity.
Moreover, beads can be separated and collected easily from the samples. For example, magnetic microbeads can be manipulated easily and controllably using a magnetic field, which strongly facilitated assay protocols.

Accordingly, magnetic beads based microfluidic platforms have been adopted by pioneering research groups studying micro scaled ChIP [59-61]. In these reported studies, antibody-functionalized beads are preloaded into microfluidic ChIP devices, which contain small valves and chambers that are fabricated in Polydimethylsiloxane (PDMS). The integrated valves permit small sample volumes of chromatin and magnetic beads to be controllably introduced and processed in miniature chambers, which can enhance antibody–target interactions and reduce incubation times. Rather than manual pipetting and multistep protocols with significant sample losses, miniature interconnecting chambers and channels manipulate samples and fluids in a repeatable, rapid manner that reduces sample consumption. For example, Quake and coworkers designed an automated microfluidic platform for ChIP that can immobilize and interrogate sonicated chromatin from as few as 2000 cells [59]. This kind of device has proved great potentials for both drug screening and antibody validation [59]. By integrating more microfluidic components in the same device, Lu and coworkers can achieve cell lysis, chromatin fragmentation, chromatin immunoprecipitation and beads purification on the same microfluidic device [60]. This kind of integration can further decrease sample loss and promote detection limitation. Coupled with real-time PCR, they characterized histone acetylation in immature 6C2 cells, using as few as 50 cells to provide biologically meaningful results within a few hours[60].
CHAPTER 2. A MICROFLUIDIC DEVICE FOR CHROMATIN IMMUNOPRECIPITATION FOLLOWED BY NEXT GENERATION SEQUENCING (ChIP-Seq) USING 100 CELLS

2.1 Introduction

Protein-DNA interaction and chromatin modification play critical roles in gene regulation. Chromatin immunoprecipitation coupled with deep sequencing (ChIP-Seq) has become the technology of choice for examining in vivo genome-wide protein-DNA interactions and chromatin modifications[62]. The assay involves covalently linking the epitope of interest to DNA by a reversible cross-linking reagent, cell lysis, immunoprecipitation of the protein of interest, reversal of the cross-linking, digestion of the protein, amplification and identification of the enriched DNA (i.e. ChIP DNA) by sequencing. A major limitation of conventional ChIP-Seq protocol is the requirement of a large number of cells (~10^7 cells). Various strategies have been developed to improve the traditional protocol over the past few years. Nano-ChIP-Seq was developed to examine histone modification using 5000 cells[37]. Single-tube linear DNA amplification (LinDA) was developed to profile the histone 3 lysine 4 tri-methylation (H3K4me3) mark using 10000 cells and oestrogen receptor-α binding using 5000 cells[38]. Both nano-ChIP-Seq and LinDA exploit novel strategies for amplifying ChIP DNA. Nano-ChIP-Seq uses a random primer with hairpin structure, optimizes conditions for faithful amplification of ChIP DNA by PCR, and uses BciVI restriction sites to allow direct ligation of Illumina sequencing adaptors[37]. LinDA amplifies ChIP DNA using an optimized T7 phage RNA polymerase linear amplification protocol that reduces amplification bias due to GC
content[38]. Besides improving amplification of ChIP DNA, the use of histone or mRNA carrier has been shown to increase recovery of ChIP DNA and allowed transcription factor ChIP-Seq using 10000 cells[63]. Indexing-first ChIP (iChIP) was recently developed to index and pool many chromatin samples before ChIP[64]. The ChIP DNA from pooled samples (containing DNA prepared from >40000 cells) was then sequenced and the data were demultiplexed based on sample-specific bar codes to yield a sensitivity of 500 cells per individual sample.

Microfluidics provides the platform for conducting molecular assays with drastic reduction in the volume, high level of integration and automation, and effective manipulation of cells and particles. Several microfluidic ChIP protocols were reported recently for studying specific loci using ChIP coupled with qPCR[36, 65, 66]. However, no effective strategies have been developed for high-efficiency collection of ChIP DNA and suppressed nonspecific adsorption at the same time. Meeting both requirements is critical for genome-wide studies (i.e. ChIP-Seq) using a small number of cells.

The sensitivity of ChIP-Seq assays is largely limited by the collection efficiency of ChIP DNA. A diploid mammalian cell contains 4-8 pg of DNA yet previous ChIP-Seq protocols could only obtain tens of picograms of DNA from 10000 cells[37, 38, 64]. Here we introduce a simple microfluidics-based protocol, microfluidic-oscillatory-washing-based ChIP-Seq (MOWChIP-Seq). It provides high collection efficiency of ChIP DNA and allows genome-wide analysis of histone modifications using as few as 100 cells. The combined use of a packed bed of beads for ChIP and effective oscillatory washing for removing nonspecific adsorption and trapping is the key to extremely high yield of highly enriched DNA.
2.2 Materials and Methods

2.2.1 Fabrication of the microfluidic ChIP device

The microfluidic chip consisted of a microfluidic chamber, connecting channels, and a micromechanical valve (Fig. 2.1 and Fig. 2.S2). The microfluidic chamber had an elliptic shape with a major axis of 6 mm, a minor axis of 3 mm and a depth of 40 µm. Micropillars were positioned inside the microfluidic chamber to prevent collapsing. The on-chip micromechanical valve, which allowed partial closure, was employed to stop magnetic IP beads while allowing liquid flow.

The microfluidic device was fabricated out of polydimethylsiloxane (PDMS) using multilayer soft lithography with minor modifications[65]. Briefly, two photomasks were generated with the microscale patterns designed using FreeHand MX (Macromedia) and printed on high-resolution (5080 dpi) transparencies. The patterns in the photomasks were replicated onto two masters (i.e. silicon wafers with photoresist patterns) for the control layer (~50 µm thick, SU-8 2025, Microchem) and the fluidic layer (~40 µm thick, SU-8 2025) with the photoresist spun on a 3-inch silicon wafer (978, University Wafer). Prepolymer PDMS (General Electric silicone RTV 615, MG chemicals) with a mass ratio of A:B = 5:1 was poured onto the fluidic layer master in a Petri dish to generate ~5 mm thick fluidic layer. PDMS at a mass ratio of A:B = 20:1 was spun onto the control layer master at 1100 rpm for 35 s, resulting in the thin PDMS control layer (~108 µm thick). Both layers of PDMS were partially cured at 80°C for 30 min. The fluidic layer was then peeled off the master. The fluidic layer feature was aligned with and bonded to that of the control layer from the top. The two-layer PDMS structure was baked at 80°C for 60 min,
peeled off from the control layer master, and punched to produce the inlet and the outlet. The two-layer PDMS and a pre-cleaned glass slide were treated with oxygen plasma cleaner (PDC-32G, Harrick Plasma) and immediately brought into contact against each other to form closed channels and chamber. Finally, the assembled chip was baked at 80°C for 1 h to strengthen the bonding between PDMS and glass. Glass slides were cleaned in a basic solution (H₂O: 27% NH₄OH: 30% H₂O₂= 5:1:1, volumetric ratio) at 75°C for 2 h and then rinsed with ultra-pure water and thoroughly blown dry.
Fig. 2. 1 Overview of the MOWChIP-Seq protocol and its optimization. (a) Schematic illustration for the five major steps of the protocol: step 1: Formation of a packed bed of IP beads; step 2: ChIP by flowing the chromatin fragments through the packed bed; step 3: Oscillatory washing; step 4: Removal of the unbound chromatin
fragments and debris by flushing the chamber; step 5: Collection of the IP beads. The microfluidic chamber contains supporting pillars (shown as small circles) that prevent collapsing. (b-d) Optimization of the MOWChIP-Seq protocol. Major parameters of the protocol were optimized by checking for IP fold enrichment of known positive (UNKL and C9orf3) and negative loci (N1 and N2). IP was done against H3K4me3 in GM12878 cells. All experiments were conducted in duplicate and the horizontal lines represent the mean. Parameters optimized include: amount of beads in device chamber (b); concentration of antibody used for coating IP beads (c); washing duration in each of the two washing buffers (d). The relative fold enrichment was normalized against that of N2. 1000-cell samples were used in (b-d). 150 µg IP beads were used in (c, d). The antibody concentration for coating was 5 µg/ml for (b) and (d). The duration of oscillatory washing was 5 min for (b, c). Flow washing in (d) was implemented by flowing each washing buffer unidirectionally for 3 min under 1.5 µl/min.

2.2.2 Setup of the microfluidic device

The microfluidic chip was mounted on an inverted microscope (IX 71, Olympus) and the operation was monitored by a CCD camera (ORCA-285, Hamamatsu) attached to the port of the microscope. Prior to experiments, the control channel was pre-filled with water to prevent bubble formation in the fluidic channel. The reagents were introduced into the inlet via perfluoroalkoxyalkane (PFA) high purity tubing (1622L, ID: 0.02 in. and OD: 0.0625 in., IDEX Health & Science) with the flow driven by a syringe pump (Fusion 400, Chemyx). The on-chip micromechanical valve was actuated by a solenoid valve.
(18801003-12V, ASCO Scientific) and a pressure source (either a gas cylinder or a compressed air outlet). A data acquisition card (NI SCB-68, National Instruments) and a LabVIEW (LabVIEW 2012, National Instruments) program were employed to control the switching of the solenoid valve. The applied pressure (35-40 psi) in the PDMS control channel deformed the thin PDMS membrane between the fluidic and control channels and closed the fluidic channel partially to stop beads while allowing liquid to flow. During oscillatory washing, the inlet and outlet of the microfluidic chamber were attached to two solenoid valves via PFA tubing and the pressure pulses were applied via the two solenoid valves under the automation by the data acquisition card and the LabVIEW program.

2.2.3 Preparation of sonicated chromatin

**10000-cell samples:** 10000-cell samples were centrifuged at 1,600xg for 5 min at room temperature in a swing bucket centrifuge with soft deceleration. Cells were then washed twice with 1.0 ml 1x PBS (14190-144, Sigma-Aldrich) at room temperature by centrifugation and resuspension. Cells were cross-linked for 5 min with 1ml 1% freshly-prepared formaldehyde (28906, Thermo Scientific). Cross-linking was terminated by adding 0.05 ml 2.5 M glycine (R000333, Covaris) and shaking for 5 min at room temperature. Cross-linked cells were pelleted and washed with pre-cooled PBS buffer and resuspended in 130 µl of the sonication buffer (Covaris, 10mM Tris-HCl, pH8.1, 1mM EDTA, 0.1% SDS, and 1× protease inhibitor cocktail (R000306, Covaris)). Cross-linked cells were sonicated with a Covaris E220 sonicator for 14 min with 5% duty cycle, 105 peak incident power and 200 cycles per burst. The sonicated lysate was centrifuged at 14000×g for 10 min at 4°C. Sonicated chromatin in the supernatant was transferred to a
new 1.5 ml LoBind Eppendorf tube (17014013, Denville) for MOWChIP-Seq. From this stock chromatin preparation, samples equivalent to 1000, 600 and 100 cells were aliquoted and diluted to give a final volume of 50 µl for MOWChIP-Seq. 10% of the sample was used as the input. After this procedure, we typically obtained ~2.7 pg DNA per cell from the pre-ChIP chromatin samples. DNA was extracted using the IPure kit from Diagenode (C03010012). DNA concentration was measured using a Qubit 2.0 fluorometer with dsDNA HS Assay kit (Q32851, Life Technologies).

100- or 600-cell samples: The procedure was different for preparing sonicated chromatin from 100 or 600 cells directly. Cells were counted with a hematocytometer and then 100 or 600 cells were transferred to a 1.5 ml LoBind Eppendorf tube containing 10 µl 10% FBS in PBS. Cells were then cross-linked for 5 min at room temperature by adding 0.625 µl 16% formaldehyde to yield a final concentration of 1%. Cross-linking was quenched by adding 1.25 µl 2.5 M glycine for 5 min at room temperature. The cross-linked sample was then diluted using 120 µl Covaris sonication buffer (to give a total volume of 130 µl) and sonicated with a Covaris E220 sonicator for 8 min with 5% duty, 105 peak incident power and 200 cycles per burst in a Covaris microtube (520045, Covaris). The sonicated lysate was centrifuged at 14000×g for 10 min at 4°C. Sonicated chromatin in the supernatant was transferred to a new 1.5 ml LoBind Eppendorf tube for MOWChIP-Seq. After this procedure, we typically obtained ~3.8 pg DNA per cell from the pre-ChIP chromatin samples. This per-cell yield was substantially higher than that obtained using the above procedure because we replaced washing of cross-linked cells.
(involving centrifugation and resuspension) with dilution by the sonication buffer to minimize chromatin loss.

### 2.2.4 Preparation of immunoprecipitation (IP) beads

Superparamagnetic Dynabeads® Protein A (2.8 μm, 30 mg/ml, 10001D, Invitrogen) were used for immunoprecipitation. 150 μg (5 μl of the original suspension) beads were washed twice with freshly-prepared IP buffer (20 mM Tris-HCl, pH8.0, 140 mM NaCl, 1mM EDTA, 0.5mM EGTA, 0.1%(w/v) sodium doxycholate, 0.1% SDS, 1%(v/v) Triton-100X) and resuspended in 150 μl IP buffer which contained antibody. Beads were gently mixed with the antibody at 4°C on a rotator mixer at 24 rpm for 2 h. Antibody-coated beads were washed twice with the IP buffer, and resuspended in 5 μl IP buffer. We optimized the antibody concentration for the bead coating step based on our ChIP-qPCR results. The optimal antibody concentration for MOWChIP-Seq with anti H3K4me3 antibody (07-473, Millipore) and anti H3K27Ac antibody (ab4729, Abcam) was 3.3 μg/ml for 100~600 cells, 5 μg/ml for 1000 cells, and 6.6 μg/ml for 10000 cells. These conditions were equivalent to using 495, 750, and 990 ng antibody in the preparation of 150 μg IP beads.

### 2.2.5 MOWChIP

The MOWChIP process involved several steps (Fig. 2.1a and Fig. 2.S2). The microfluidic device was first rinsed with the IP buffer for conditioning. The antibody-coated magnetic IP beads were then loaded into the microfluidic chamber via the combined effects of pressure-driven flow (provided by the syringe pump) and magnetic force...
generated by a cylindrical permanent magnet (NdFeB, D48-N52, 0.25 in. dia. and 0.5 in. thick, K&J Magnetics). The on-chip micromechanical valve was partially closed and the IP beads were packed against the valve to form a packed bed. After the loading of the IP beads (~150 µg under optimal condition), the IP buffer (with freshly added 1 mM PMSF (78830-1G, Sigma-Aldrich) and 1% protease inhibitor cocktail (P8340, Sigma-Aldrich)) containing sonicated chromatin fragments (with a total volume of either 50 or 130 µl) was flowed through the packed bed of IP beads at a flow rate of 1.5 or 3.5 µl/min, respectively. Under these flow rates, the immunoprecipitation step was finished around 40 min.

After ChIP, a low-salt washing buffer (20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 2 mM EDTA, 0.1%SDS, 1%(v/v) Triton-100X) was flowed into the microfluidic chamber. Oscillatory washing was conducted (for 5 min unless otherwise noted) to remove nonspecifically adsorbed or physically trapped materials from the bead surface. We prefilled the tubing with 10 µl washing buffer at each end of the microfluidic chamber and kept the on-chip valve open. Pressure pulses (each at 3 psi, with a pulse width of 0.5 s and an interval of 0.5 s between two pulses) were applied alternatingly at either end of the microfluidic chamber. The duration and frequency of the pressure pulses were set in a LabVIEW program and implemented via the regulation of the two solenoid valves by the data acquisition card (Fig. 2.S1b). After the oscillatory movement, the IP beads were retained by the NdFeB magnet on one side of the chamber while the unbound chromatin fragments and other debris/waste were flushed out of the microfluidic chamber by a clean washing buffer flow at 2 µl/min. The process of oscillatory washing was repeated once using a high salt washing buffer (20 mM Tris-HCl, pH 8.0, 500 mM NaCl, 2 mM EDTA, 0.1% SDS, 1%(v/v) Triton-100X). Finally the IP beads were flowed out of the microfluidic
chamber under a flow rate of 50 μl/min and collected into a 1.5 ml LoBind Eppendorf tube containing 100 μl IP buffer. The optimal duration for washing was 5 min for both washing buffers.

**Fig. 2.S 1** Setup of the microfluidic device with ancillary control and liquid delivery equipment. Each microfluidic device has one inlet, one outlet, and one port for...
applying pressure to the control channel to actuate the on-chip pneumatic microvalve via its connection with a solenoid valve. (a) The inlet is connected to a syringe pump when a reagent is delivered (while the outlet is open to the atmosphere). (b) During oscillatory washing, the inlet and outlet are attached to two separate solenoid valves that are alternatingly actuated. (c) The microfluidic chip on a microscope stage. There are 6 parallel microfluidic devices on the same chip.
Fig. 2.S 2 Microscopic images that recorded the MOWChIP process. The scale bar represents 1 mm. (a) The microfluidic chamber is filled with buffer. (b) The pressurization of the control channel partially closes the pneumatic microvalve (i.e. liquid passes the microvalve while beads get stopped). (c) The IP beads are stopped by the partially closed
microvalve and form a packed bed. (d) Oscillatory washing by applying alternating pressure pulses on the inlet and the outlet after the flow of the chromatin fragments through the packed bed. (e) The retention of IP beads using a magnet (not shown) while flushing out undesired molecules and debris.

2.2.6 Extraction of ChIP DNA and input DNA

Chromatin samples (either ChIP or input chromatin) were processed by IPure kit (C03010012, Diagenode) to extract DNA. Purified DNA was dissolved in 10 μl DNase-free water and used directly for ChIP-qPCR or for sequencing library construction. DNA concentrations were measured using a Qubit 2.0 fluorometer with dsDNA HS Assay kit (Q32851, Life Technologies).

2.2.7 Construction of sequencing libraries

Sequencing libraries were prepared using ThruPLEX-FD kit (Rubicon Genomics). This kit reduces the assay time and the risk of contamination by using a single tube and eliminating intermediate purification steps. The process involved template preparation, library synthesis, and library amplification. Adaptor-based PCR amplification (98 °C for 20 s, 72 °C for 50 s for each cycle) was used during library amplification. We used 11 cycles for input DNA, 12~13 cycles for ChIP DNA from 10000 cells, and 14~17 cycles for ChIP DNA from 1000 or fewer cells. The libraries were purified using Ampure XP beads (A63880, Beckman Coulter). Library fragment size was determined using high sensitivity DNA analysis kit (5067-4626, Agilent) on an Agilent 2200 TapeStation. KAPA library
quantification kit (KK4809, Kapa Biosystems) was used to determine effective library concentrations. The final concentrations of libraries submitted for sequencing were ~2 nM. The libraries were sequenced on an Illumina HiSeq 2500 with single-end 50 nt read. Typically 15-20 million reads were generated per library.

2.2.8 Cell culture

GM12878 cells were obtained from Coriel Institute for Medical Research. Species of origin of the cell line was confirmed by PCR targeting the glucose-6-phosphate dehydrogenase gene. Donor subject has a single bp (G-to-A) transition at nucleotide 681 in exon 5 of the CYP2C19 gene (CYP2C19*2) which creates an aberrant splice site. Donor origin of the cell line was confirmed using PCR against the point mutation. The cell line was tested for mycoplasma contamination using ABI MycoSEQ mycoplasma detection assay (Applied Biosystems). Cells were propagated in RPMI 1640 (11875-093, Gibco) plus 15% fetal bovine serum (26140-079, Gibco), 100 U penicillin (15140-122, Gibco), and 100mg/ml streptomycin (15140-122, Gibco) at 37°C in a humidified incubator containing 5% CO₂. Cells were sub-cultured every two days to maintain them in exponential growth phase.

2.2.9 Mouse strain, embryo dissection and cell sorting by fluorescence-activated cell sorting (FACS)

The University of Iowa Office of the Institutional Animal Care and Use Committee review board approved these studies. Wild type C57BL/6 (Stock No. 000664) and
B6129S6F1 (Stock No. 101043) mice were purchased from the Jackson Laboratory. To obtain embryonic day 14.5 (E14.5) fetal liver (FL), B6129SF1 females were mated with C57BL/6 males (6-9 weeks old) late in the afternoon and females were checked the following morning for the presence of a vaginal plug, which was designated as E0.5. FLs were dissected from E14.5 embryos. Single cell suspensions were prepared by dissociating mechanically and expelling the cells through 40 µm nylon filter (352340, Falcon), followed by red blood cell lysis (ACK Lysing Buffer, 10-548E, Lonza). Cells were resuspended in 1 ml staining buffer (2% FBS in PBS) per $1 \times 10^8$ cells. FACS was performed as previously described[67, 68] with a few modifications. To remove nonspecific binding, anti-mouse CD16/CD32 (Fc Block, 101302, Biolegend) were added to single cell suspension and incubated for 10 min at 4°C. Next, cells were stained with a cocktail of antibodies against lineage markers (Ly-6G/Ly-6C (108417), CD45R/B220 (103225), CD3ε (100321), TER-119b (116215), CD4 (100529), CD8a (100723), CD19 (115521)), Kit (17-1171-83), and Sca-1 (12-5981-83). Lineage antibodies were purchased from Biolegend. Kit and Sca-1 antibodies were purchased from eBiosciences. Stained samples were first subjected to yield sort for Lineage-Sca-1+ Kit+ (LSK) cells and collected into a 12×75-mm polystyrene tube containing 500µl 1× IMDM (12440-053, Gibco)+20% FBS. Collected cells were then subjected to purity sort using the same gating strategy and sorted into a 1.5 ml DNA LoBind tube containing 0.8 ml 1× IMDM+50% FBS. On average, ~10000 FL LSK cells can be obtained per mouse embryo.

2.2.10 MOWChIP-qPCR data analysis
Real-time PCR was done using iQ SYBR Green Supermix (1708882, Bio-Rad) on a CFX96 real-time PCR machine (Bio-Rad) with C1000Tm thermal cycler base. All PCR assays were performed using the following thermal cycling profile: 95°C for 10 min followed by 40 cycles of (95°C for 15 s, 58°C for 40 s, 72°C for 30 s). Primer concentrations were 400 nM. All primers were ordered from Integrated DNA Technologies. The results were represented as relative fold enrichment, which is the ratio of percent input between a positive locus and a negative locus. Percent input was computed using the following equation:

\[
Percent \ input = 100 \times 2^{(Ct_{\text{input}} - \log(DF) \div \log(2) - Ct_{\text{IP}})}
\]

where \(Ct_{\text{input}}\) and \(Ct_{\text{IP}}\) are the Ct values of input and ChIP DNA, respectively; dilution factor (DF) is defined as (sample volume of input + sample volume of IP)/(sample volume of input).

### 2.2.11 MOWChIP-Seq reads mapping and normalization

Sequencing reads were mapped to the mouse genome (mm9) and human genome (hg19) using Bowtie2 (v2.2.2) [69] with default parameter settings. Uniquely mapped reads from both ChIP and input samples were used to compute a normalized signal for each 100 nt bin across the genome. Normalized signal is defined as following:

\[
\text{normalized signals} = \text{IP}\left(\frac{\text{Reads in each bin}}{\text{Total uniquely mapped reads}} \times 1000000\right) - \\
\text{Input}\left(\frac{\text{Reads in each bin}}{\text{Total uniquely mapped reads}} \times 1000000\right).
\]
2.2.12 Peak calling of MOWChIP-Seq data

Only uniquely mapped reads were used for peak calling. Two peak callers were used with the following parameter settings: MACS (p-value < 10^{-5}) [70] and SPP (z-score > 4) [71] with other parameters set at default values. The final set of high-confidence peaks was those that were called by both methods.

2.2.13 Construction of Receiver Operating Characteristic (ROC) curves

Using ROC curves, we compared the performance of MOWChIP-Seq to that of two state-of-the-art methods nano-ChIP-seq[37] and iChip[64] . We focused on promoter regions (defined as 2000 bp upstream and 500 bp downstream of a transcription start site (TSS)). We obtained published ChIP-Seq data generated using conventional protocol with a large sample size (typically 10 million cells per sample) as the gold standard. The gold-standard true positives were defined as the set of high-confidence promoter peaks identified as described in the peak calling section. The set of promoter regions that did not overlap with any peaks were defined as the gold-standard negative set. Using the gold-standard sets, the following quantities were defined to compute the ROC curve: True Positives (TPs), peaks that were supported by the gold-standard positive set; False Positives (FPs), peaks that were not supported by the gold-standard positive set; False Negatives (FNs), gold-standard positives that were not called peaks in an experiment; True Negatives (TN), peaks that were not called in an experiment and were in the gold-standard negative set. True positive rate (TPR) was defined as TP/(TP+FN) and false
positive rate (FPR) was defined as FP/(FP+TN). ROC curves were generated by computing TPR and FPR values on prediction sets obtained by varying the peak calling threshold.

The gold-standard datasets used for constructing the ROC curves for MOWChIP-Seq, nano-ChIP-seq and iChIP were summarized in Table 2.2. Briefly, we generated H3K4me3 and H3K27ac data using GM12878 cells. The corresponding gold-standard data were generated by the ENCODE consortium. The authors of Nano-ChIP-Seq generated H3K4me3 data using mouse ESCs. The corresponding gold-standard data were from Marson et al.[72] and Goren et al.[73]. The authors of iChIP generated H3K4me3 data using mouse CD4 T and B cells. The corresponding gold-standard data were from Wei et al.[74] and Heinz et al.[75].

2.2.14 Correlation analysis of MOWChIP-Seq data with other published ChIP-Seq data sets

To evaluate the quality of our FL HSPC data, we selected four published datasets of H3K4me3 and H3K27ac using BM LSK[37], BM LT-HSC[64], B cell[64] and macrophage[64]. For a given histone modification, normalized ChIP-Seq signals in all promoter regions in the genome were extracted. Promoter regions were defined as +/-2kb around transcription start sites (TSS). TSS annotation was based on RefSeq. Averaged signals across the promoter region was used. Promoter regions with zero signals in both data sets were excluded for computing Pearson correlation coefficient.

2.2.15 Prediction of enhancers and super enhancers using epigenomic data
We used “H3K4me3lo + H3K27Achih” to define enhancers. Specifically, enhancers were predicted using the CSI-ANN algorithm[76] and normalized H3K4me3 and H3K27ac MOWChIP-Seq signals across the genome. The coordinates of the predicted enhancers and H3K27ac MOWChIP-Seq data were then used as the input to predict super enhancers using the ROSE software by the Young lab (http://bitbucket.org/young_computation/rose). We set the parameters to allow enhancers within 15000 bp to be stitched together. In addition, we excluded the constituent enhancers located within +/- 2000 bp from annotated TSSs.

2.2.16 Transcription factor motif enrichment analysis

We compiled a set of 1207 TF binding motifs from three major public databases, JASPAR [77], UniPROBE [78], Transfac [79] and motifs of ten hematopoietic transcription factors[80]. We used the program CentriMo [81] to identify over-represented motifs in a given set of enhancer sequences. Default parameters of CentriMo were used.

2.2.17 Gene ontology (GO) term enrichment analysis of super enhancer targets

Genes closest to the super enhancers were used as their targets. Database for Annotation, Visualization, and Integrated Discovery (DAVID) [82] was used for GO analysis of the target genes. Nominal p-values were corrected for multiple testing using the method by Benjamini and Hochberg [83]. GO terms with a corrected p-value of 0.05 were regarded as significant.
2.2.18 Assumptions of statistical tests

All statistical tests were performed using large sample sizes and underlying distribution assumptions were met. Sample sizes were reported in Fig. legends. All reported p-values were corrected for multiple testing.

2.3 Results and Discussion

2.3.1 Design and Operation of MOWChIP

We used multilayer soft lithography to design and fabricate a polydimethylsiloxane (PDMS) device, featuring a simple microfluidic chamber (~710 nl in volume) for high-efficiency ChIP. The microfluidic chamber has one inlet 1 and one outlet 2 and the outlet has an on-chip pneumatic microvalve that can be partially closed by exerting a pressure at 3[65, 84] (Fig. 2.1a, Fig. 2.S1 and Fig. 2.S2). First, magnetic beads (~2.8 µm in diameter and coated with a ChIP antibody) are flowed into the microfluidic chamber and form a packed bed while the pneumatic microvalve is partially closed. Sonicated chromatin fragments (~200-600 bp) are then flowed through the packed bed of IP beads and adsorbed onto the bead surface. When closely packed, the gaps among the IP beads are smaller than 2 µm and facilitate rapid and high-efficiency adsorption of targeted chromatin fragments under the small diffusion length. The IP beads are then washed by oscillatory washing in two different washing buffers to remove nonspecifically adsorbed chromatin fragments. Finally, the IP beads (with adsorbed chromatin fragments) are flowed out of the chamber and collected for off-chip processing. The entire on-chip process takes ~1.5 h.
2.3.2 Optimization of MOWChIP

We found that the quality and amount of ChIP DNA were affected by several parameters of the protocol, including amount of IP beads in the device, antibody concentration used for coating IP beads, duration of the oscillatory washing, and cell sample size. We optimized these parameters by using MOWChIP-qPCR to examine fold enrichment at known positive and negative loci for H3K4me3 (with the primer sequences listed in Table 2.1) in a human lymphoblastoid cell line, GM12878 (Fig.2.1b-d).

| GM12878 and H3K4me3 | C9orf3 | F       | CCTCCTCAGTTCTCCAGACT |
|                     |        | R       | AGCTGAGGTGGTAAGATGTGAC |
|                     | UNKL   | F       | CAGCCACCCACCTAGGAA    |
|                     |        | R       | TCCTATGGCTCCCCAGGT    |
|                     | N1     | F       | TCATCTGCAAATGGGGACAA  |
|                     |        | R       | AGGACACCCCCTCTCAACAC  |
|                     | N2     | F       | ATGGTTGCCACTGGGGATCT  |
|                     |        | R       | TGCCAAAGCCTAGGGGAAGA  |

*Table 2.1 Primer sequences used in MOWChIP-qPCR.*

The fold enrichment reached a peak value at an intermediate bead amount, likely due to increased nonspecific adsorption and trapping when too many beads were used (Fig. 2.1b). Similarly, for antibody used for coating IP beads, we obtained the highest fold enrichment at an intermediate antibody concentration (Fig.2. 1c). This is likely due to insufficient antibody coverage on the beads at low concentration, which decreases the amount of binding for chromatin targets. On the other hand, excessive amount of antibody
on the beads may promote binding to low-affinity or nonspecific chromatin. The high-efficiency adsorption by the packed bed of IP beads also led to increased nonspecific adsorption and physical trapping. We found that oscillatory washing was essential for the high quality of ChIP DNA (Fig. 2.1d). At the same time, excessive washing needs to be avoided in order to reduce DNA loss. Using optimized conditions that balanced both DNA yield and quality, we were able to obtain \(~1.3\) ng ChIP DNA from 10000 cells (5.3\% of the total chromatin) and \(~180\) pg from 1000 cells (6.2\% of the total chromatin) for H3K4me3 after DNA purification (Fig. 2.S3). This yield was almost 2 orders of magnitude higher than that reported in previous work\[37\] and within the range of the theoretical limit (2.2-7.8\% of the genome is marked by H3K4me3 based on ENCODE data\[85\]). To assess the amount of background reads in our data, we computed the fraction of reads in peaks (FRiP)\[86\]. The values were 35.6\% and 21.6\% for our 10000- and 1000-cell data, respectively. These were substantially higher than 1\% guideline recommended by ENCODE, suggesting low background in our recovered chromatin. As a result, we used ChIP DNA directly for sequencing library construction without pre-amplification.
Fig. 2.S 3  The yield of ChIP DNA as a function of antibody concentrations used during coating of the IP beads. ChIP was done using anti-H3K4me3 antibody and GM12878 cells. The chromatin samples were processed by IPure kit (C03010012, Diagenode) to extract DNA before the measurement. The concentrations of DNA in these samples were measured by a Qubit 2.0 fluorometer using dsDNA HS Assay kit (Q32851, Life Technologies). Data points are mean of two replicate experiments. Error bars represent s.d.. Under the optimized MOWChIP-Seq conditions (i.e. antibody concentration of 5 µg/ml for 1000-cell samples and 6.6 µg/ml for 10000-cell samples), 1000 and 10000 cells yielded 180 pg and 1.3 ng ChIP DNA, respectively. In comparison, the same numbers of cells yielded 2.9 ng and 24.6 ng DNA without the ChIP step.

2.3.3 Quantification of MOWChIP-seq using cell line cells
To evaluate the performance of MOWChIP-Seq, we used it to profile H3K4me3 and histone 3 lysine 27 acetylation (H3K27Ac) marks with various amounts of chromatin from GM12878 cells. We prepared sonicated chromatin using 10000 cells and aliquoted chromatin samples equivalent to 10000, 1000, 600 and 100 cells. For all four cell sample sizes, replicate experiments were highly correlated (average $r = 0.933$ and 0.894 for H3K4me3 and H3K27Ac, respectively, Fig. 2.S4). Using published ChIP-Seq data generated by conventional protocol with millions of cells per sample as the gold standard, we compared the performance of MOWChIP-Seq to two other methods, nano-ChIP-Seq[37] and iChIP[64] (detailed in Materials and Methods and Table 2.2).

<table>
<thead>
<tr>
<th>Histone Modification</th>
<th>GEO Accession Number</th>
<th>Reference (PMID number)</th>
<th>Cell Type</th>
<th>Number of cells</th>
<th>Purpose in this study</th>
</tr>
</thead>
<tbody>
<tr>
<td>H3K4me3, H3K27Ac</td>
<td>GSE29611 GSE35583</td>
<td>ENCODE GM12878</td>
<td>millions</td>
<td>ROC curve construction</td>
<td></td>
</tr>
<tr>
<td>H3K4me3</td>
<td>GSE22075</td>
<td>Adli et al. (20622861)</td>
<td>ESC, BM LSK</td>
<td>20000</td>
<td>ROC curve construction (ESC data) and genome-wide correlation (BM LSK data)</td>
</tr>
<tr>
<td>H3K4me3</td>
<td>GSE11724</td>
<td>Marson et al. (18692474)</td>
<td>ESC</td>
<td>millions</td>
<td>ROC curve construction</td>
</tr>
<tr>
<td>H3K4me3</td>
<td>GSE12241</td>
<td>Goren et al. (19946276)</td>
<td>ESC</td>
<td>mililons</td>
<td>ROC curve construction</td>
</tr>
</tbody>
</table>
### Table 2. Published ChIP-Seq data sets used in this study.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Accession</th>
<th>Authors</th>
<th>Cell Type</th>
<th>Sample Size</th>
<th>ROC curve construction</th>
</tr>
</thead>
<tbody>
<tr>
<td>H3K4me3, H3K27Ac</td>
<td>GSE60103</td>
<td>Lara-Astiaso et al.</td>
<td>BM LT-HSC, CD4 T cell, B cell, macrophage</td>
<td>5000</td>
<td>(CD4 T and B cells) and genome-wide correlation (BM LT-HSC, B, and macrophage)</td>
</tr>
<tr>
<td>H3K4me3</td>
<td>GSE14254</td>
<td>Wei et al. (19144320)</td>
<td>CD4 T cell</td>
<td>millions</td>
<td>ROC curve construction</td>
</tr>
<tr>
<td>H3K4me3</td>
<td>GSE21512</td>
<td>Heinz et al. (20513432)</td>
<td>B cell</td>
<td>millions</td>
<td>ROC curve construction</td>
</tr>
</tbody>
</table>

We used the Receiver Operating Characteristic (ROC) curve to quantify the data quality in terms of the agreement with a gold standard. The Area Under the ROC Curve (AUC) is a standard metric for quantifying balanced sensitivity and specificity. For H3K4me3, MOWChIP-Seq with 100-600 cells showed performance that was comparable to iChIP with 5000 cells and superior to nano-ChIP-Seq with 20000 cells (Fig. 2.2a). For H3K27Ac, MOWChIP-Seq also produced data with excellent AUC values when using as few as 100 cells (Fig.2.2b). Normalized MOWChIP-Seq signals (see Online Methods) for H3K4me3 at the Spi1 gene locus show consistency among samples of various sizes (Fig. 2.2c). This gene encodes an important transcription factor for B cell development. The Spi1 promoter region was highly enriched for H3K4me3 signal in MOWChIP-Seq data generated using 100-10000 cells. H3K27Ac is a mark for active transcriptional enhancers[87]. We show normalized MOWChIP-Seq signals for H3K27Ac across the
immunoglobin heavy chain locus (Fig.2.2d). As expected, the AUC values decreased with decreasing number of cells. Nevertheless, all data had good quality and reproducibility that enabled analysis of important genome-wide features (Table 2.3).

**Fig. 2.2** MOWChIP-Seq generates high quality data using as few as 100 cells. The performance of MOWChIP-Seq was compared to those of two other methods: nano-ChIP-seq and iChIP. (a) Receiver Operating Characteristic (ROC) curves for H3K4me3 data. ROC curves were constructed by comparing the ChIP-Seq data generated by various methods to published gold-standard data generated using conventional protocols with millions of cells. Nano-ChIP-Seq data was from Adli et al. [37];
iChIP data was from Lara-Astiaso et al. [64]. Values shown are average Area Under the ROC curve (AUC) of two replicate experiments. (b) ROC curves for H3K27Ac data generated by MOWChIP-Seq. (c) Normalized H3K4me3 MOWChIP-Seq signals at the SPI1 gene locus using data generated with various sample sizes. ENCODE data were generated using millions of cells and shown for comparison. (d) Normalized H3K27Ac MOWChIP-Seq signals at the immunoglobulin heavy chain locus. Known B-cell enhancers are indicated at the bottom of the Fig.
**Fig. 2.S 4** Genome-wide correlations among MOWChIP-Seq data generated using various GM12878 cell sample sizes. Normalized ChIP-Seq signals from promoter regions across the genome were used for computing correlations (n = 34342). Promoter regions were defined as +/- 2kb around transcription start sites (TSS). TSS annotation was based on RefSeq. Promoter regions with zero values in both samples were removed before computing correlation coefficients. Colors represent Pearson correlation coefficients. (a) H3K4me3. (b) H3K27Ac.

<table>
<thead>
<tr>
<th># Cells</th>
<th>Total num. of reads (Million)</th>
<th>Total num. of mapped reads (Million)</th>
<th>% duplicate reads</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>GM12878 H3K4me3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10K</td>
<td>20.2</td>
<td>19.4</td>
<td>79</td>
<td>0.978</td>
</tr>
<tr>
<td></td>
<td>19.2</td>
<td>18.7</td>
<td>64</td>
<td>0.979</td>
</tr>
<tr>
<td>1K</td>
<td>21.3</td>
<td>19.9</td>
<td>89</td>
<td>0.977</td>
</tr>
<tr>
<td></td>
<td>20.7</td>
<td>18.5</td>
<td>76</td>
<td>0.978</td>
</tr>
<tr>
<td>600</td>
<td>11.3</td>
<td>10.9</td>
<td>58</td>
<td>0.975</td>
</tr>
<tr>
<td></td>
<td>16.6</td>
<td>15.9</td>
<td>72</td>
<td>0.978</td>
</tr>
<tr>
<td>100</td>
<td>15.3</td>
<td>12.6</td>
<td>80</td>
<td>0.946</td>
</tr>
<tr>
<td></td>
<td>11.9</td>
<td>11.1</td>
<td>81</td>
<td>0.92</td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
<td>------</td>
<td>----</td>
<td>------</td>
</tr>
<tr>
<td><strong>GM12878 H3K27Ac</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10K</td>
<td>16.2</td>
<td>15.5</td>
<td>76</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>19.4</td>
<td>19.2</td>
<td>79</td>
<td>0.98</td>
</tr>
<tr>
<td>1K</td>
<td>17.2</td>
<td>16.1</td>
<td>74</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>15.5</td>
<td>15.0</td>
<td>77</td>
<td>0.97</td>
</tr>
<tr>
<td>600</td>
<td>12.3</td>
<td>11.3</td>
<td>72</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>18.1</td>
<td>17.1</td>
<td>75</td>
<td>0.97</td>
</tr>
<tr>
<td>100</td>
<td>10.9</td>
<td>10.3</td>
<td>71</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>15.3</td>
<td>14.3</td>
<td>71</td>
<td>0.92</td>
</tr>
<tr>
<td><strong>FL LSK H3K4me3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10K</td>
<td>20.6</td>
<td>19.4</td>
<td>55</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>13.9</td>
<td>13.2</td>
<td>55</td>
<td>NA</td>
</tr>
<tr>
<td>1K</td>
<td>11.4</td>
<td>10.5</td>
<td>60</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>10.9</td>
<td>9.1</td>
<td>67</td>
<td>NA</td>
</tr>
<tr>
<td>600</td>
<td>13.2</td>
<td>9.0</td>
<td>50</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>9.6</td>
<td>8.2</td>
<td>66</td>
<td>NA</td>
</tr>
<tr>
<td>100</td>
<td>17.1</td>
<td>13.7</td>
<td>58</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>52.2</td>
<td>45.8</td>
<td>64</td>
<td>NA</td>
</tr>
</tbody>
</table>
Table 2.3 Summary statistics of MOWChIP-Seq data generated in this study. All mapped reads were used to compute the percentage of duplicate reads.

2.3.4 MOWChIP-seq of HSPCs

We applied MOWChIP-Seq to study the epigenome of hematopoietic stem and progenitor cells (HSPCs) isolated from mouse fetal liver. Little is known about the dynamics of the epigenome during embryonic hematopoiesis, largely due to the difficulty in isolating sufficient quantities of highly purified HSCs from developing embryos. This challenge makes mouse HSPC study an ideal test case for our technology. Definitive HSCs first appear in embryonic day 10.5 aorta-gonads-mesonephros and thereafter migrate to the fetal liver (FL) where they proliferate before they eventually colonize bone marrow (BM). Previously, histone modifications had been mapped in adult BM HSCs but not stem and progenitor cells at any earlier stage. We mapped H3K4me3 and H3K27Ac using chromatin equivalent to 10000, 1000, 600 and 100 purified FL HSPCs (Fig. 2.S5).
For all four cell sample sizes, replicate experiments were highly correlated (average $r = 0.864$ and 0.881 for H3K4me3 and H3K27Ac, respectively, Fig. 2.S6). We also computed the correlation of our data with three published datasets on BM HSPC data. Since the HSPCs being compared are highly related but not identical, our data showed lower but reasonable correlation with published data, further supporting the quality of our data (average $r = 0.677$ and 0.745 for H3K4me3 and H3K27Ac, respectively, Fig.2.S6). Histone modification signals of the promoter regions were also correlated with gene expression levels ($r$ ranged from 0.47 to 0.60, Fig. 2.S7). Taken together, these results suggest that our FL HSPC data is of high quality.
Fig. 2.5 Purification of hematopoietic stem and progenitor cells (HSPCs) from mouse fetal liver. HSPCs were purified as Lineage- Sca-1+ Kit+ (LSK) cells using fluorescence activated cell sorting. Cells were double sorted with a yield sort first followed by a purity sort. Percentages shown in each panel indicate the purity of fetal liver LSK cells.
Fig. 2.S 6 Genome-wide correlations between MOWChIP-Seq data on FL HSPCs data and published data. Normalized ChIP-Seq signals from promoter regions across the genome were used for computing correlations (n = 34342). Promoter regions were defined as +/- 2kb around transcription start sites (TSS). TSS annotation was based on RefSeq. Promoter regions with zero values in both samples were removed before computing correlation coefficients. Colors represent Pearson correlation coefficients.
Bone marrow (BM) LSK data was obtained from Adli et al. BM LT-HSC, B cell and macrophage data were obtained from Lara-Astiaso et al. (a) H3K4me3. (b) H3K27Ac.

Fig. 2.S 7 Scatter plots of histone modification signals and transcript expression levels.

FL LSK microarray data was downloaded from the GEO database (accession no.: GSE21404). All transcripts annotated in the RefSeq database were used for computing Spearman correlation coefficient (n = 24658). X-axis, normalized expression values of
Little is known about the enhancer repertoire of FL HSPCs. Using our MOWChIP-Seq data, we predicted active enhancers in FL HSPCs using the signature of H3K4me3\textsuperscript{lo} + H3K27Ac\textsuperscript{hi}. In total, we predicted 10407, 6523, 7083, and 6909 enhancers (FDR < 0.5%) using the 10000, 1000, 600, and 100-cell data sets, respectively. The average pairwise overlap of the four sets of enhancers was 81.8% (Fig. 2.3a). In total, 4446 enhancers were shared among all four data sets, which we used as the final set of enhancers in this study (Fig. 2.3a). We identified many known transcriptional enhancers that are active in FL HSPCs, such as the Tal1\textsuperscript{+19} enhancer (Fig. 2.3b) \cite{88}, Erg\textsuperscript{+85} enhancer and Runx1\textsuperscript{+24} enhancer (Fig. 2.S8) \cite{89, 90}. DNA motif analysis revealed that the set of enhancers were enriched for binding motifs of 45 transcription factors, including many well-known hematopoietic TFs such as ERG, ETV6, FLI1, PU.1 and RUNX1. To identify the unique enhancers in FL HSPCs, we compared our enhancer set to the enhancer catalog covering 16 blood cell types\cite{64}. We found that 58% (2,561) of enhancers identified in this study were unique to FL HSPCs, suggesting enhancer activity is highly dynamic during early hematopoiesis.
Fig. 2. 3 Epigenomics-aided discovery of novel enhancers and super enhancers in fetal liver HSPCs. (a) Venn diagram of sets of enhancers predicted using epigenomic data generated using various sample sizes. (b) Normalized H3K27Ac ChIP-Seq signals at the known Tal +19 enhancer. BM_HSC denotes data on bone marrow HSC generated by Lara-Astiaso et al. [64]. (c) Venn diagram of sets of super enhancers predicted using epigenomic data generated using various sample sizes. (d) Normalized H3K27Ac ChIP-Seq signals at the super enhancer of the Flt3 gene that plays a critical role in hematopoiesis.
**Fig. 2.S** Normalized H3K27Ac MOWChIP-Seq signals at known enhancers in FL HSPCs. Data generated using different numbers of cells are shown. (a) Erg +85 enhancer. (b) Runx1 +24 enhancer. BM_HSC data was obtained from Lara-Astiaso et al.
Super enhancer (SE) is a newly discovered class of enhancers that are typically much longer than single enhancers [91]. They play a critical role in regulating genes that determine lineage identity. Almost nothing is known about super enhancers in HSPCs. Using our epigenomic data, we discovered 131 SEs in FL HSPCs (Fig. 2.3c). Consistent with the notion that SEs often regulate lineage-conferring genes, target genes of our predicted SEs were enriched for genes involved in hematopoiesis (p-value = 6.5E-3, hypergeometric test). Example target genes included many known key regulators of hematopoiesis such as Erg, Etv6, Fli1, Flt3, Runx1, and Spi1. The super enhancer controlling the Flt3 gene plays an important role in hematopoiesis, especially for FL HSPCs (Fig. 2.3d)[92].

2.3.5 MOWChIP-seq using small number of cells

The chromatin used for generating 1000-, 600-, and 100-cell data above (Fig. 2.2 and 2.3) was aliquoted from a stock chromatin sample prepared from 10000 cells. In order to use MOWChIP-Seq directly on samples with 100-600 cells, we replaced washing of cross-linked cells with dilution by the sonication buffer in order to minimize chromatin loss due to centrifugation and resuspension (see Online Methods). Our modified cross-linking and sonication procedures generated desired chromatin size distribution for sequencing after library preparation (Fig. 2.S9). We generated additional MOWChIP-Seq data using the modified protocol with starting cell numbers of 100 and 600 and assessed the data quality by correlation and ROC curve analysis. Data generated using the two protocols (i.e. using the stock chromatin vs. starting directly with 100 or 600 cells) were highly correlated (average r = 0.843 over two histone marks, Fig.2.S10). Data generated using the two protocols also had similar AUC values (Fig. 2.2 and Fig. 2.S11).
Fig. 2. S 9 DNA fragment size distribution after sonication and library preparation with starting cell numbers of 600 and 100. DNA size distribution was examined by an Agilent 2100 Bioanalyzer using high sensitivity DNA analysis kit (5067-4626, Agilent). (a) DNA size distribution using 600 cells as the starting material and before library preparation. DNA was purified after pre-ChIP processing and MOWChIP. (b) DNA size distribution using 600 cells as the starting material and after library preparation. DNA was purified after pre-ChIP processing, MOWChIP, and library preparation. (c) DNA size distribution using 100 cells as the starting material and after library preparation. DNA was purified after pre-ChIP processing, MOWChIP, and library preparation.
Fig. 2.S10 Genome-wide correlations between MOWChIP-Seq data generated using two protocols for chromatin preparation. Chromatin was prepared from GM12878 cells. “From stock chromatin”, cross-linked and sonicated chromatin was prepared from 10000 cells. Chromatin amounts equivalent to 100 and 600 cells were then aliquoted from the stock chromatin preparation. “From cells directly”, chromatin was prepared directly from 100 and 600 starting cells. Values shown are Pearson correlation coefficients. Normalized ChIP-Seq signals from promoter regions across the genome were used for computing correlations (n = 50961). Promoter regions were defined as +/- 2kb around transcription start sites (TSS). TSS annotation was based on RefSeq. Promoter regions with zero values in both samples were removed before computing correlation coefficients. Duplicate experiments were done using each protocol. (a) H3K4me3 data using 600 cells. (b) H3K4me3 data using 100 cells; (c) H3K27Ac data using 600 cells. (d) H3K27Ac data using 100 cells.

<table>
<thead>
<tr>
<th></th>
<th>From Stock Chromatin</th>
<th>From Cells Directly</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. H3K4me3 (600)</td>
<td>0.938 0.948</td>
<td>0.940 0.946</td>
</tr>
<tr>
<td></td>
<td>0.818 0.837</td>
<td>0.801 0.825</td>
</tr>
<tr>
<td>b. H3K4me3 (100)</td>
<td>0.879 0.877</td>
<td>0.884 0.886</td>
</tr>
<tr>
<td></td>
<td>0.763 0.765</td>
<td>0.679 0.700</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>From Stock Chromatin</th>
<th>From Cells Directly</th>
</tr>
</thead>
<tbody>
<tr>
<td>c. H3K27Ac (600)</td>
<td>0.879 0.877</td>
<td>0.884 0.886</td>
</tr>
<tr>
<td></td>
<td>0.763 0.765</td>
<td>0.679 0.700</td>
</tr>
<tr>
<td>d. H3K27Ac (100)</td>
<td>0.879 0.877</td>
<td>0.884 0.886</td>
</tr>
<tr>
<td></td>
<td>0.763 0.765</td>
<td>0.679 0.700</td>
</tr>
</tbody>
</table>
**Fig. 2.** MOWChIP-Seq data using 100 and 600 starting cells. Chromatin was prepared from GM12878 cells. (a) and (b), Receiver Operating Characteristic (ROC) curves. ROC curves were constructed by comparing MOWChIP-Seq data to published gold-standard data generated using conventional protocols with millions of cells. Values shown are average Area Under the ROC curve (AUC) of two replicate experiments. (a)
H3K4me3 data. (b) H3K27Ac data. (c) and (d), Normalized MOWChIP-Seq signals at the SPI1 gene locus (c) and at the immunoglobin heavy chain locus (d). (c) H3K4me3 data. (d) H3K27Ac data.

2.4 Conclusions

In summary, we demonstrated that MOWChIP-Seq with as few as 100 cells generated high quality genome-wide profiles of histone modifications. Our microfluidic technology is fundamentally different from other high-sensitivity ChIP technologies which rely on superior amplification[37, 38, 93] and indexing-pooling[64] schemes, thus may potentially complement other methods. The microfluidic device has a simple structure and is easy to operate. The platform allows running of multiple assays in parallel. Our technology paves the way for epigenomic studies involving extremely low number of cells from animals and patients.

Accession codes. Gene Expression Omnibus: MOWChIP-Seq data were deposited under accession number GSE65516.
3.1 Introduction

The phenotype of a cell is determined by its gene expression profile and its response to environmental cues. It has long been recognized that gene activity is not only affected by DNA sequence, but also by mechanisms that affect gene expression and cellular phenotypes without altering DNA sequence (i.e. epigenetics). Cellular processes are regulated by a complex interplay among different layers of epigenetic information, including DNA methylation, histone modifications, nucleosome positions, and expression of noncoding RNA[216]. DNA methylation is the best studied epigenetic modification in mammals. DNA methylation typically refers to the addition of a methyl group at the carbon-5 position of cytosine residues within CpG dinucleotides, forming 5-methylcytosine (5mC). This process is catalyzed by enzymes called DNA methyltransferases. Clusters of CpG sites (“CpG islands” or CGIs) commonly span promoters of housekeeping genes. Promoter CGIs typically remain unmethylated in normal cells and are associated with active gene expression during differentiation, while methylated CGIs are associated with gene repression. The genome-wide DNA methylation profile (i.e. the DNA methylome) is established early in development for regulation and maintenance of gene expression during differentiation. The methylome is altered and disrupted in disease states. For example, transcriptional silencing of tumor-suppressor genes by CGI-promoter hypermethylation plays an important role in cancer development[217-219]. A better understanding of the dynamics in the
methylome during disease development will improve various aspects of biomedicine, including risk stratification, disease diagnostics, and drug/therapeutic discovery.

DNA methylation is critically involved in cancer development. The hypermethylation at various genes in tumors has been recognized as a common epigenetic feature for all types of human cancers. Genes that were critically involved in cancer biology, including the cell-cycle inhibitor p16-INK4a and the DNA-repair genes MLH1 and BRCA1, have been shown to undergo methylation-associated silencing in tumor cells[220, 221]. While acquiring hypermethylation at specific promoters, overall human tumors undergo a massive loss of DNA methylation. Since DNA methylation protects the integrity of chromosomes, this global DNA hypomethylation potentially contributes to the large-scale genetic changes that are hallmarks of tumorigenesis.

Profiling DNA methylomes contributes to various aspects of cancer intervention in the context of personalized medicine. First, methylomic profiles may serve as highly-sensitive cancer cell markers because CGI hypermethylation of tumor-suppressor genes occurs early in the tumorigenesis and the number of genes undergoing this epigenetic alteration increases during cancer development. Second, methylomic profiles can provide valuable information for prognosis because certain patterns of methylation-based silencing are associated with tumor aggressivity and angiogenesis[222]. Third, methylomes have also been used to predict responses to a therapeutic procedure or reagent[223]. Finally, methylomic marks have been used as therapeutic targets. DNA-demethylating agents have been explored as promising drugs for reversing promoter DNA hypermethylation and associated gene silencing[224].
There have been a number of genome-wide technologies developed for profiling DNA methylomes. Bisulfite sequencing is generally considered the gold standard for DNA methylation analyses[225, 226]. Bisulfite treatment converts cytosine residues to uracil, but leaves 5mC residues unaffected. Such treatment introduces changes in the DNA sequence based on the methylation status of individual cytosine residues. Combined with next-generation sequencing (NGS), the approach generates methylomic profile with single-nucleotide resolution. In spite of the ultrahigh resolution, bisulfite sequencing requires a large amount of DNA (1-5 µg) and the cost associated with deep sequencing (>500 million reads per sample to cover ~95% of the genome) can be prohibitive when various samples with repeats are analyzed. Enrichment-based technologies have also been applied to study methylomes. These approaches dramatically reduce the required sequencing depth by enriching methylated DNA fragments based on affinity purification. As a prime example, methylated DNA immunoprecipitation followed by next-generation sequencing (MeDIP-Seq) uses a monoclonal antibody specific for 5mC to target single-stranded methylated DNA fragments (sheared by sonication) and identifies these fragments using sequencing (shown in Fig. 3.1 below)[227-230]. In comparison to bisulfite sequencing, MeDIP requires only 30-60 million reads per sample with a 100-300 bp resolution. MeDIP-seq provides an unbiased tool for probing DNA methylation events both within CGIs and non-CGI regions throughout the entire genome.

DNA methylomes, similar to other epigenetic information, are specific to cell and tissue types, the disease condition and its developmental stage. Profiling methylomes associated with various tissues and diseases is critical for understanding the dynamics in
methylomes during disease development and establishing epigenomic signatures for
disease diagnosis and prognosis. Such goal is actively pursued in efforts such as
Roadmap Epigenomics Program (a consortium funded by NIH common fund). However,
one critical challenge in these efforts is that current methylomic tools often do not offer
sufficient sensitivity to examine tiny quantities of cell samples from scarce sources such
as small lab animals and patients. For example, conventional whole-genome bisulfite
sequencing requires 1-5 µg DNA and MeDIP-seq requires 5-20 µg DNA (due to the low
efficiency associated with immunoprecipitation). Given that a diploid mammalian cell
typically contains 4-8 pg of DNA, bisulfite sequencing and MeDIP-seq require at least 10^6-
10^7 cells. In contrast, mouse and patient samples do not yield large quantities of cells.
For example, the number of naturally occurring T regulatory cells in murine splenocytes
is ~ 10,000 per spleen, and ~5000 per ml peripheral blood leukocyte. Circulating tumor
cells (CTCs) are present by the frequency of 1-10 per ml of whole blood in patients with
metastatic cancer. Furthermore, the isolation of a homogenous single cell type always
generates further loss in the sample amount. Thus highly sensitive technologies for
methylome profiling are in high demand, in order to facilitate generation of data with direct
biomedical relevance.
In this project, we present an ultrasensitive microfluidic MeDIP-seq technology for profiling methylomes. Based on our preliminary study, we believe that microfluidic MeDIP-seq with high immunoprecipitation efficiency will produce a sensitivity of 0.1 ng DNA (or ~50-100 cells). This will be roughly 4-5 orders of magnitude higher than the prevailing protocol and 2-3 orders of magnitude higher than the state-of-the-art (~50 ng)[231]. As a proof of concept, we will use the technology to obtain methylomic profiles in a transgenic mouse model during mammary cancer development. Our data will generate insights into the temporal dynamics in the methylome during the disease process.
3.2 Materials and methods

3.2.1 Fabrication of the microfluidic MeDIP device

The microfluidic chip consisted of a microfluidic chamber, connecting channels, and a micromechanical valve. The microfluidic chamber had an elliptic shape with a major axis of 6 mm, a minor axis of 3 mm and a depth of 40 μm. Micropillars were positioned inside the microfluidic chamber to prevent collapsing. The on-chip micromechanical valve, which allowed partial closure, was employed to stop magnetic beads while allowing liquid flow.

The microfluidic device was fabricated out of polydimethylsiloxane (PDMS) using multilayer soft lithography with minor modifications. Briefly, two photomasks were generated with the microscale patterns designed using FreeHand MX (Macromedia) and printed on high-resolution (5080 dpi) transparencies. The patterns in the photomasks were replicated onto two masters (i.e. silicon wafers with photoresist patterns) for the control layer (~50 μm thick, SU-8 2025, Microchem) and the fluidic layer (~40 μm thick, SU-8 2025) with the photoresist spun on a 3-inch silicon wafer (978, University Wafer). Prepolymer PDMS (General Electric silicone RTV 615, MG chemicals) with a mass ratio of A:B = 5:1 was poured onto the fluidic layer master in a Petri dish to generate ~5 mm thick fluidic layer. PDMS at a mass ratio of A:B = 20:1 was spun onto the control layer master at 1100 rpm for 35 s, resulting in the thin PDMS control layer (~108 μm thick). Both layers of PDMS were partially cured at 80°C for 30 min. The fluidic layer was then peeled off the master. The fluidic layer feature was aligned with and bonded to that of the control layer from the top. The two-layer PDMS structure was baked at 80°C for 60 min, peeled off from the control layer master, and punched to produce the inlet and the outlet.
The two-layer PDMS and a pre-cleaned glass slide were treated with oxygen plasma cleaner (PDC-32G, Harrick Plasma) and immediately brought into contact against each other to form closed channels and chamber. Finally, the assembled chip was baked at 80°C for 1 h to strengthen the bonding between PDMS and glass. Glass slides were cleaned in a basic solution (H₂O: 27% NH₄OH: 30% H₂O₂ = 5:1:1, volumetric ratio) at 75°C for 2 h and then rinsed with ultra-pure water and thoroughly blown dry.

### 3.2.2 Setup of the microfluidic device

The microfluidic chip was mounted on an inverted microscope (IX 71, Olympus) and the operation was monitored by a CCD camera (ORCA-285, Hamamatsu) attached to the port of the microscope. Prior to experiments, the control channel was pre-filled with water to prevent bubble formation in the fluidic channel. The reagents were introduced into the inlet via perfluoroalkoxyalkane (PFA) high purity tubing (1622L, ID: 0.02 in. and OD: 0.0625 in., IDEX Health & Science) with the flow driven by a syringe pump (Fusion 400, Chemyx). The on-chip micromechanical valve was actuated by a solenoid valve (18801003-12V, ASCO Scientific) and a pressure source (either a gas cylinder or a compressed air outlet). A data acquisition card (NI SCB-68, National Instruments) and a LabVIEW (LabVIEW 2012, National Instruments) program were employed to control the switching of the solenoid valve. The applied pressure (35-40 psi) in the PDMS control channel deformed the thin PDMS membrane between the fluidic and control channels and closed the fluidic channel partially to stop beads while allowing liquid to flow. During oscillatory washing, the inlet and outlet of the microfluidic chamber were attached to two
solenoid valves via PFA tubing and the pressure pulses were applied via the two solenoid valves under the automation by the data acquisition card and the LabVIEW program.

3.2.3 Preparation of DNA sample

Genomic DNA was extracted from $10^6$ cell samples using Blood & Cell Culture DNA Mini Kit (Qiagen). Genomic DNA was extracted from mouse tissue using DNeasy Blood & Tissue Kit (Qiagen). Extracted genomic DNA were sonicated with a Covaris E220 sonicator for 180s with 10% duty cycle, 50 peak incident power and 200 cycles per burst. The concentration of sonicated DNA was quantified using a Qubit 2.0 fluorometer with dsDNA HS Assay kit (Q32851, Life Technologies). Different sample sizes (e.g. 0.5~10ng) were aliquoted and diluted (with MeDIP buffer) to give a final volume of 50 µl for MeDIP. 10ng of the sonicated sample was used as the input. Sonicated DNA samples were freshly denatured into single strand DNA (ssDNA) under 97°C for 15 min and put on ice before the IP.

3.2.4 Preparation of immunoprecipitation (IP) beads

Superparamagnetic Dynabeads® Protein A (2.8 µm, 30 mg/ml, 10001D, Invitrogen) were used for immunoprecipitation. 150 µg (5 µl of the original suspension) beads were washed twice with freshly-prepared MeDIP buffer (10 mM monobasic sodium phosphate dihydrate, 10mM dibasic sodium phosphate, 140 mM NaCl, 0.05%(v/v) Triton-100X) and resuspended in 150 µl MeDIP buffer which contained antibody. Beads were gently mixed with the antibody at 4°C on a rotator mixer at 24 rpm for 2 h. Antibody-coated beads were
washed twice with the MeDIP buffer, and resuspended in 5 µl MeDIP buffer. The antibody concentration for MeDIP-seq with 5-Methylcytosine (5-mC) antibody (pAb) (61255, Active Motif) was 3.3 µg/ml for 0.5ng DNA, 5 µg/ml for 1ng DNA, and 6.6 µg/ml for 10ng DNA. These conditions were equivalent to using 495, 750, and 990 ng antibody in the preparation of 150 µg IP beads.

3.2.5 Microfluidic MeDIP

The MeDIP process involved several steps. The microfluidic device was first rinsed with the MeDIP buffer for conditioning. The antibody-coated magnetic beads were then loaded into the microfluidic chamber via the combined effects of pressure-driven flow (provided by the syringe pump) and magnetic force generated by a cylindrical permanent magnet (NdFeB, D48-N52, 0.25 in. dia. and 0.5 in. thick, K&J Magnetics). The on-chip micromechanical valve was partially closed and the IP beads were packed against the valve to form a packed bed. After the loading of the beads (~150 µg under optimal condition), the MeDIP buffer containing denatured DNA fragments (with a total volume of 50 µl) was flowed through the packed bed of IP beads at a flow rate of 1.5 µl/min, respectively. Under these flow rates, the immunoprecipitation step was finished around 40 min.

After IP, fresh MeDIP buffer was flowed into the microfluidic chamber. Oscillatory washing was conducted (for 5 min unless otherwise noted) to remove nonspecifically adsorbed or physically trapped materials from the bead surface. We prefilled the tubing with 10 µl washing buffer at each end of the microfluidic chamber and kept the on-chip valve open. Pressure pulses (each at 3 psi, with a pulse width of 0.5 s and an interval of 0.5 s between two pulses) were applied alternately at either end of the microfluidic
chamber. The duration and frequency of the pressure pulses were set in a LabVIEW program and implemented via the regulation of the two solenoid valves by the data acquisition card. After the oscillatory movement, the beads were retained by the NdFeB magnet on one side of the chamber while the unbound chromatin fragments and other debris/waste were flushed out of the microfluidic chamber by a clean washing buffer flow at 2 µl/min. The process of oscillatory washing was repeated once using fresh MeDIP buffer. Finally the beads were collected out of the microfluidic chamber using a magnet.

3.2.6 Elution of MeDIP DNA

Magnetic beads after MeDIP were processed by IPure kit (C03010012, Diagenode) to elute DNA. Purified DNA was dissolved in 10 µl DNase-free water and used directly for sequencing library construction.

3.2.7 Construction of sequencing libraries

Sequencing libraries were prepared using DNA SMART™ ChIP-Seq kit (Clontech). This highly sensitive kit generates robust sequencing libraries from low-input ssDNA MeDIP samples(100pg~10ng), with minimal handling, in about four hours. The libraries were purified using Ampure XP beads (A63880, Beckman Coulter). Library fragment size was determined using high sensitivity DNA analysis kit (5067-4626, Agilent) on an Agilent 2200 TapeStation. KAPA library quantification kit (KK4809, Kapa Biosystems) was used to determine effective library concentrations. The final concentrations of libraries submitted for sequencing were ~2 nM. The libraries were sequenced on an Illumina HiSeq 2500 with single-end 50 nt read. Typically 15-20 million reads were generated per library.
3.2.8 Cell culture

GM12878 cells were obtained from Coriel Institute for Medical Research. Species of origin of the cell line was confirmed by PCR targeting the glucose-6-phosphate dehydrogenase gene. Donor subject has a single bp (G-to-A) transition at nucleotide 681 in exon 5 of the CYP2C19 gene (CYP2C19*2) which creates an aberrant splice site. Donor origin of the cell line was confirmed using PCR against the point mutation. The cell line was tested for mycoplasma contamination using ABI MycoSEQ mycoplasma detection assay (Applied Biosystems). Cells were propagated in RPMI 1640 (11875-093, Gibco) plus 15% fetal bovine serum (26140-079, Gibco), 100 U penicillin (15140-122, Gibco), and 100mg/ml streptomycin (15140-122, Gibco) at 37°C in a humidified incubator containing 5% CO₂. Cells were sub-cultured every two days to maintain them in exponential growth phase.

3.2.9 Analysis of sequencing data

MeDIP sequencing reads were mapped to the human genome (hg19) using Bowtie2 (v2.2.5) with default parameter settings. Peaks of each MEDIP sample were called against input by MACS (v1.4.2) with (p-value < 10⁻⁵) and other parameters set at default values. Pearson correlations of genome-wide coverage profiles between samples were calculated by Bioconductor R package "MEDIPS" with MEDIPS.correlation function.
3.3 Results and discussion

In our this project, we tested a simple microfluidic device for ultrasensitive MeDIP-seq. We used multilayer soft lithography to design and fabricate a polydimethylsiloxane (PDMS) device, featuring a single microfluidic chamber (~710 nl in volume) for high-efficiency IP. The microfluidic chamber has one inlet 1 and one outlet 2 and the outlet has an on-chip pneumatic microvalve that can be partially closed by exerting a pressure at 3[65]. Key features and operation of the device are shown in Fig. 3.2. First, magnetic IP beads (~2.8 µm in diameter and coated with 5-Methylcytosine (5-mC) antibody (pAb) from Active Motif) are flowed into the microfluidic chamber and form a packed bed while the pneumatic microvalve is partially closed. Single-stranded DNA fragments (~200-600 bp, produced by extracting genomic DNA from GM12878 cells using Blood & Cell Culture DNA Mini Kit from Qiagen, shearing the DNA in a Covaris E220 sonicator, and denaturing to generate ssDNA by heating at 97°C for 15 min) are then flowed through the packed bed of IP beads and adsorbed onto the bead surface. The fragment size of sample DNA, which can be quantified using gel electrophoresis or bioanalyzer(Fig.6.3), is important for the success of the IP. When closely packed, the gaps among the IP beads are smaller than 2 µm and facilitate rapid and high-efficiency adsorption of targeted DNA fragments under the small diffusion length. The IP beads are then washed by oscillatory washing[232] in a buffer to remove nonspecifically adsorbed DNA fragments. During oscillatory washing, pressure pulses (each at 3 psi, with a pulse width of 0.5 s and an interval of 0.5 s between two pulses) are applied alternatingly at either end of the microfluidic chamber. The duration and frequency of the pressure pulses are set in a LabVIEW program and implemented via the regulation.
of the two solenoid valves by a data acquisition card. After the oscillatory movement, the beads are retained by a NdFeB magnet on one side of the chamber while the unbound DNA fragments and other debris/waste are flushed out of the microfluidic chamber by a clean washing buffer. Finally, the IP beads (with adsorbed DNA fragments) are flowed out of the chamber and collected for off-chip processing. The entire on-chip process takes ~1.5 h. We found that the combined use of a packed bed of beads for MeDIP and effective oscillatory washing for removing nonspecific adsorption/trapping is the key to extremely high yield of highly-enriched MeDIP DNA.

Fig. 3. 2 Schematic illustration for the five major steps of microfluidic MeDIP. step 1: Formation of a packed bed of IP beads; step 2: MeDIP by flowing the DNA fragments through the packed bed; step 3: Oscillatory washing; step 4: Removal of the unbound DNA fragments and debris by flushing the chamber with buffer; step 5: Collection of the IP beads. The microfluidic chamber contains supporting pillars (shown as small circles) that prevent collapsing.
Fig. 3. 3 DNA fragment size distribution after sonication. DNA size distribution was examined by an Agilent 2200 TapeStation using high sensitivity DNA analysis kit (5067-4626, Agilent).

The preliminary MeDIP-seq data using genomic DNA obtained from GM12878 cells is shown in Fig.6.4. The results that we obtained using our microfluidic technology and with various amounts of starting DNA (100-0.5 ng). As expected, the quality of methylomic profiles declined when the DNA amount decreased from 100 to 0.5 ng (Fig. 3.4a). Under these unoptimized conditions, our technique was able to generate high-quality methylomic profiles using as little as 5 ng DNA. The correlation coefficient r between the two replicates was excellent for samples of 100-5 ng (0.93, 0.91, 0.90 for 100-, 10-, 5-ng samples, respectively) (Fig. 3.4b). Using the high-quality 100-ng samples as the gold standard, we got an average r of 0.91 and 0.90 for 10- and 5-ng samples’ correlations to the gold standard. For 0.5 ng samples, the correlation between the two replicates was 0.80 and the correlations to the gold standard were 0.69 and 0.60 for the two replicates. Nevertheless, most of the major peaks can be seen in the 0.5-ng data (Fig.
3. 4a). Obviously there is a lot of room for technological improvement in the sub-1 ng range.

![Image of preliminary microfluidic MeDIP-seq data]

**Fig. 3.4 Preliminary microfluidic MeDIP-seq data.** (a) Genome browser tracks for our microfluidic MeDIP-seq using various amounts of DNA (100-0.5 ng). Each sample size was tested in duplicate. (b) Genome-wide correlations among MeDIP-seq data sets of various sample sizes. Normalized MeDIP-seq signals across the genome were used for computing correlations. Colors represent Pearson correlation coefficients.

### 3.4 Conclusions

In summary, we demonstrated that MeDIP-seq with tiny amount of DNA (<5ng) generated high quality genome-wide profiles of DNA methylation modifications. Our microfluidic technology may potentially complement other methods. The microfluidic device has a simple structure and is easy to operate. The platform allows running of multiple assays in parallel. Our technology paves the way for epigenomic studies involving extremely low number of cells or DNA from animals and patients.
4.1 Introduction

Epigenetics is defined as heritable changes in gene activity and expression that occur without alteration in DNA sequence[19]. It is known these non-genetic alternations are tightly regulated by two major epigenetic modifications: histone proteins associated with DNA (histone modifications) and chemical modifications to the cytosine residues of DNA (DNA methylation)[19]. Immunoprecipitation of target chromatin (ChIP) or methylated DNA (MeDIP) with specific epigenetic mark has become the technology of choice for examining epigenetic information[94, 95].

However, conventional ChIP or MeDIP assay requires a large number of cells (10^6~10^7 cells) or DNA (1µg~10 µg ) as starting material. This is largely due to inevitable sample loss, potential sample contamination or denaturing during conventional sonication and immunoprecipitation steps, which involves tedious manual handling steps including multiple tube transferring, centrifugation, pipetting and vortexing. This limitation poses a serious challenge for studying rare or limiting samples during the processes of stem cell differentiation, embryo development and oncogenesis. Conventional ChIP/MeDIP is also time consuming due to inefficient immunoprecipitation. Typically, overnight incubation is necessary for the immunoprecipitation step and it takes several days to complete the whole procedure.
Microfluidics provides the platform for conducting molecular assays with drastic reduction in the volume, high level of integration and automation, and effective manipulation of cells and molecules. Several microfluidic ChIP protocols were reported recently. Some of the protocols adopted enzymatic digestion to achieve on-chip chromatin fragmentation directly from cells. However, enzymatic fragmentation of chromatin or DNA suffers from biases of chromatin region or sequence-specific cleavage. In most of the protocols, chromatin was first fragmented by using off-chip sonication and then loaded into the microfluidic device for ChIP reaction. Off-chip sonication avoids sequence bias but inevitably leads to sample loss due to multiple tube transferring. In addition, conventional sonicators are not able to manipulate samples with volume less than 100µl. These combined factors seriously restrict the potential of microfluidics for probing small amount or small volume of samples. Moreover, off-chip sonication using conventional sonicators increases the risk of sample contamination (especially with probe type sonicator). Last but not least, sonication generates a large amount of heat, which lead to sample degradation or denaturing. The equipped cooling system of conventional sonicators (e.g. Bioruptor or Covaris) only reflect the ambient temperature, i.e. in the water bath, and fail to monitor the real temperature inside samples. Therefore, potential overheating during sonication could be unnoticed. Because of the noted problems associated with off-chip chromatin fragmentation, previously reported microfluidic assay did not achieve fully integrated ChIP assay starting directly from cells. In addition, previous microfluidic ChIP assays rely on either simple diffusion or valve-actuated mixing to achieve immunoprecipitation (IP), which normally took ~2h. A pure diffusion-based mixing process is inefficient and time-consuming, particularly when the solution...
contained macromoleculars, such as DNA or chromatin, that have diffusion coefficient orders of magnitude lower than most of liquid. The configuration of valve-based device increases the complexity of on-chip structure and requires cumbersome controlling system including computer programmed solenoid valve, tubing, compressed air, etc.

In order to address the noted limitations, here we describe a novel microfluidic device integrating a composite piezoelectric transducer for both on-chip sonication and acoustofluidic enhanced IP in highly controlled way. In our device, the piezoelectric transducer was actuated at its resonant frequency and the generated lamb wave (a type of transverse wave) was transferred into the neighboring micro-chamber. With high amplitude, the acoustic field was employed to shear cross-linked cells for chromatin releasing and chromatin/DNA fragmentation into a wide and controllable range of fragment sizes. With low amplitude, the acoustic field was able to generate acoustic streaming to enhance micro-scale mixing during IP and the following washing steps inside the micro-chamber. The combination of both sonication and IP in the same device can decrease the sample loss/contamination due to tube transferring and significantly decrease the assay time as well due to microscale reaction and acoustic enhanced mixing. Using this device, we demonstrated microfluidic ChIP/MeDIP assay starting from as few as ~100 cross-linked cells or 500 pg of genomic DNA that were finished within ~30 min. As a proof-of-concept, we used this protocol to examine histone modification H3K4me3 and DNA methylation (5-mC) of GM 12878 cells, which is a lymphoblastoid cell line. The patterns of noted histone modification and DNA methylation has been well-documented in previous work. The ability of our system to perform ChIP/MeDIP on low cell number
and with dramatically reduce assay time was demonstrated by comparing relative fold enrichment at known positive locus against known negative locus using qPCR analysis.

4.2 Methods and materials

4.2.1 Fabrication of microfluidic device

The microfluidic device is advantageously fabricated out of polydimethylsiloxane (PDMS) using previously described soft lithography techniques [98]. Briefly, two photomasks were first generated with the microscale patterns designed by computer-aided design software FreeHand MX (Macromedia, San Francisco, CA, USA) and printed on high-resolution (5,080 dpi) transparencies. The master were made of negative photoresist SU-8 2075 (~100µm of depth, Microchem, Newton, MA, USA) by spin-coating on a 3-inch silicon wafer (University Wafer, South Boston, MA, USA) at corresponding speed. Afterwards, PDMS at a mass ratio of RTV615 A: RTV615 B = 10 : 1 was poured onto the master in a Petri dish to generate ~5 mm thick PDMS layer. The PDMS layer was solidified by baking at 80 °C for ~120 min, peeled off from the flow layer master, and punched to produce inlet and outlet reservoirs. Then the PDMS layer and a pre-cleaned glass slide were treated with oxygen plasma and immediately brought into contact against the slide to form sealed micro-chamber. Finally, the assembled chip was baked at 80°C for another 1 h to promote the bonding strength between PDMS and glass. Glass slides were pre-cleaned in a basic solution (H2O: 27% NH4OH: 30% H2O2= 5 : 1 : 1, volumetric ratio) at 75°C for 2 h and then rinsed with ultrapure water and thoroughly blown dry.
4.2.2 Biological sample preparation

GM12878 cells were obtained from Coriel Institute for Medical Research. Species of origin of the cell line was confirmed by PCR targeting the glucose-6-phosphate dehydrogenase gene. Donor subject has a single bp (G-to-A) transition at nucleotide 681 in exon 5 of the CYP2C19 gene (CYP2C19*2) which creates an aberrant splice site. Donor origin of the cell line was confirmed using PCR against the point mutation. The cell line was tested for mycoplasma contamination using ABI MycoSEQ mycoplasma detection assay (Applied Biosystems). Cells were propagated in RPMI 1640 (11875-093, Gibco) plus 15% fetal bovine serum (26140-079, Gibco), 100 U penicillin (15140-122, Gibco), and 100mg/ml streptomycin (15140-122, Gibco) at 37°C in a humidified incubator containing 5% CO₂. Cells were sub-cultured every two days to maintain them in exponential growth phase.

The harvested cells were centrifuged at 300g for 5 min and resuspended in culture medium. The concentration of cells in the medium was obtained using a hemocytometer. Designated number of cells, ranging from 100~100k, were cross-linked in 160 µl culture medium with 1% formaldehyde (28908, Thermo Scientific) for 5 min and the crosslinking was then terminated by adding 2.5M freshly-made glycine (final concentration is 0.125 M) and incubating for 5 min at room temperature. The cross-linked cells were then pelleted and resuspended in 7 µl of ChIP IP buffer (20 mM Tris-HCl, pH8.0, 140 mM NaCl, 1mM EDTA, 0.5mM EGTA, 0.1%(w/v) sodium doxycholate, 0.1% SDS, 1%(v/v) Triton-100X, with freshly added 1 mM PMSF (78830-1G, Sigma-Aldrich) and 1% protease inhibitor cocktail (P8340, Sigma-Aldrich)) and ready to be loaded into the microfluidic chamber.
Genomic DNA (gDNA) of GM 12878 cells was extracted and purified using QIAamp DNA blood mini kit (51104, Qiagen). The extracted genomic DNA was dissolved in water to a final concentration of 50ng/µl before use. Designated amount of gDNA, ranging from 500pg to 50ng, was diluted in 7ul of MeDIP IP buffer and incubated under 97°C for 15 min. The denatured DNA was ready to be loaded into the microfluidic chamber.

4.2.3 Setup of the sonication integrated microfluidic ChIP/MeDIP (SIM-ChIP/MeDIP) device

A schematic of the SIM-ChIP device is shown in Fig. 4.1. A Langevin type transducer (MPI-2525D-60H, UltrasoncWorld) was mounted to the glass slide of the microfluidic chip using two-component epoxy glue (2-Ton Epoxy, Devcon) cured overnight. A function generator was used to generate sinusoidal AC signal, which was amplified by an high-power ultrasonic actuator (PDUS200, Micromechatronics Inc.). The amplified AC signal could be switched on/off by a Labview programmed relay (5501, Coto Technology, North Kingstown, RI) via a data requisition card (NI SCB-68, National Instruments, Austin, TX). The signal was employed to drive the transducer at its resonant frequency. The resulting working voltage and current were monitored through a build-in monitor port on the amplifier via a digital oscilloscope (B&K precision2530, Newark, Chicago, IL, USA). The resonance frequency of the device was slightly shifted from the fundamental resonance frequency of the transducer (60KHz) due to bonding to the microfluidic device. The working resonance frequency (59~63 KHz) can be pinpointed by tracking the frequency which yield the maximum output current. To monitor the temperature inside the micro-chamber, a cement-on surface thermalcouple (CO1-T, Omega Engineering Inc.,
Stamford, CT) was attached to the bottom surface of the microfluidic device, just beneath the micro-chamber. The detected temperature can be displayed and recorded by a temperature controller (CN8202, Omega Engineering Inc.). The whole device set was put on an ice bag, which could facilitate the cooling during the sonication. For microscopic observation, the device was mounted on an inverted microscope (IX 71, Olympus) and the operation was monitored by a CCD camera (ORCA-285, Hamamatsu) attached to the port of the microscope.

4.2.4 Preparation of immunoprecipitation (IP) beads

Superparamagnetic Dynabeads® Protein A (2.8 μm, 30 mg/ml, 10001D, Invitrogen) were used for immunoprecipitation. 150 µg (5 µl of the original suspension) beads were washed twice with freshly-prepared IP buffer and resuspended in 150 µl IP buffer which contained antibody. Beads were gently mixed with the 6 µg/ml antibody at 4°C on a rotator mixer at 24 rpm for 1 h. Antibody-coated beads were washed twice with the IP buffer, and resuspended in 2 µl IP buffer.

4.2.5 SIM-ChIP/MeDIP

Fig. 4.4 shows the SIM-ChIP procedure. First of all, the crosslinked cells or gDNA, which were suspended in freshly-made IP buffer, were loaded into the micro-chamber using a pipette via the inlet. After cells/gDNA loading, the inlet and outlet was tightly sealed using an adhesive and transparent sealing tape (Microseal “B” Adhaesive seals, Biorad, Hercules, CA.) to avoid solution loss during sonication. Then the transducer was driven at its resonance frequency with high acoustic intensity (V=25V, and I=0.5A) to perform on-chip sonication for 5 min. During this process, the computer programmed
relay was able to automatically switch on/off the transducer and break single long sonication period into multiple short cycles with rest time in between. Each short cycle has 10 seconds, with 3 second on and 7 second off.

After on-chip sonication, the antibody pre-coated magnetic IP beads were then loaded into the microfluidic chamber via the magnetic force generated by a cylindrical permanent magnet (NdFeB, D48-N52, 0.25 in. dia. and 0.5 in. thick, K&J Magnetics). After beads loading, the transducer was actuated at its resonance frequency with low acoustic intensity (with V=10V and I=0.2A) to perform acoustic-enhanced on-chip IP. At this amplitude, the beads slurry and sonicated sample solution could be fully mixed by acoustic streaming in less than 1 min. The acoustic-enhanced on-chip IP was finished in 30 min.

After on-chip IP, the magnetic IP beads were retained inside the micro-chamber using the magnet, while the sample solution was replaced by introducing the washing buffer into the micro-chamber. For ChIP experiment, a low-salt washing buffer (20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 2 mM EDTA, 0.1% SDS, 1%(v/v) Triton-100X) was flowed into the microfluidic chamber. For MeDIP experiment, MeDIP washing buffer was loaded into the chamber. The transducer was actuated again with low acoustic intensity (V=10V and I=0.2A) to perform acoustic-enhanced on-chip washing, which took about 2 min. The on-chip washing was repeated once using a high-salt washing buffer (20 mM Tris-HCl, pH 8.0, 500 mM NaCl, 2 mM EDTA, 0.1% SDS, 1%(v/v) Triton-100X) for ChIP experiment, and using the same MeDIP washing buffer for MeDIP experiment. Finally the IP beads were collected out of the microfluidic chamber into a 0.6 ml LoBind Eppendorf tube containing 100 µl IP buffer using both magnet and pipette.
For reference input DNA, we specifically on-chip sonicated same number of cells or same amount of gDNA, which did not undergone any IP, under exactly the same sonication condition as ChIP/MeDIP samples.

4.2.6 Extraction of ChIP/MeDIP DNA and input DNA

ChIP/MeDIP samples, as well as the input samples, were purified using Ipure kit (C03010012, Diagenode, Inc.). Purified DNA was dissolved in 10 µl DNase-free water and used directly for ChIP-qPCR. DNA concentrations were measured using a Qubit 2.0 fluorometer with either dsDNA HS assay kit (Q32851) or ssDNA HS Assay kit (Q10212, Life Technologies).

4.2.7 Sample fragment size analysis using gel electrophoresis

Sonicated chromatin was first reverse-crosslinked with 100 µL elution buffer containing 1% SDS and 0.5 mg/mL proteinase K, and then incubated at 65 °C for 2 h. De-crosslinked chromatin was subsequently extracted using Phenol-chloroform extraction and precipitated using ethanol precipitation. The purified DNA was finally eluted in 10 µL water and analyzed on a 1% agarose gel. Sonicated gDNA was analyzed directly on an GelRed (41003, Biotium, Inc.) stained 1% agarose gel. The size of the band was determined using100 bp ~ 3 kb DNA ladders (100bp PCR molecular ruler, 170-8206, Bio-rad,Inc.).

4.2.8 Real-time PCR
Real-time PCR was done using iQ SYBR Green Supermix (1708882, Bio-Rad) on an CFX96 real-time PCR machine (Bio-Rad) with C1000Tm thermal cycler base. All PCR assays were performed using the following thermal cycling profile: 95°C for 10 min followed by 40 cycles of (95°C for 15 s, 58°C for 40 s, 72°C for 30s). Primer concentrations were 400 nM. All primers were ordered from Integrated DNA Technologies. The ChIP-qPCR results were represented as relative fold enrichment, which is the ratio of percent input between a positive locus and a negative locus. Percent input was computed using the following equation:

$$\text{Percent input} = 100 \times 2^{\frac{\text{Ct}_{\text{input}} - \log(\text{DF})}{\log 2} - \text{Ct}_{\text{IP}}}$$

where Ct_{input} and Ct_{IP} are the Ct values of input and ChIPed DNA, respectively; dilution factor (DF) is defined as (sample volume of input + sample volume of IP)/(sample volume of input).

### 4.3 Results and discussion

#### 4.3.1 Design of the microfluidic sonication device

The microfluidic sonication system, as shown in Fig. 4.1, consists several major parts: a function generator generated the driving AC signal, which was amplified by an signal amplifier. The amplified signal actuated the bolt clamped Langevin transducer bound on a microfluidic chip at its resonance frequency (~60 kHz). The Langevin transducer, which has a structure of piezoelectric elements “sandwiched” between two pieces of metal, can generate much higher acoustic power than that generated by regular piezo transducers[99, 100]. Such high acoustic power carried by the transverse acoustic wave
was transferred into the micro-chamber. The microfluidic chamber, as shown in Fig. 4.1, has simple one-layer structure. It had a diameter of 8 mm and a height of 100 mm. The inlet and outlet were punctured inside the chamber, as shown in Fig. 4.1, to infuse sample solution into the chamber.

After the actuation of the transducer, solution in the micro-chamber was then subjected to rapid changes of pressure. If the amplitude of pressure change is beyond certain threshold, small air bubbles, which were originally dissolved inside the liquid, could rapidly collapse. This process is referred as inertial cavitation, which was reported to be the main cause of chromatin/DNA fragmentation during sonication[101, 102]. Because cavitation initiates at the gas-liquid interface, the intensity of cavitation was promoted by increasing the gas-liquid interface via micro-fabricated “crescent” shaped structures, as shown in Fig. 4.1. Acoustic cavitation comprises several stages: nucleation, growth and collapse. The “crescent” PDMS structures can enhance the cavitation in several ways. First of all, the “crescent” structures entrapped air inside the microfluidic chamber (as shown in Fig. 4.1), thus increase the air-liquid interface, where most of the cavitation took place[103]. Secondly, the hydrophobicity of PDMS material also promote the cavitation phenomenon. Previous studies indicate that air phase presented at the liquid and hydrophobic solid interface could be involved in the cavitation process as they have a size close to critical radii (in a range of sub-micrometer) for nucleation[104-107]. Therefore, by creating micro-fabricated PDMS structures inside the chamber, we increased the interface between liquid and hydrophobic solid interface, thus increased the potential locations where cavitation can happen. Thirdly, as shown in Fig. 4.2b, the sharp edge of “crescent” structure (with a tip angle about 12°) can significantly promote
the initiation of cavitation via increased acoustic oscillation at the tip of the sharp edge, because the vibrational amplitude of the tip, as well as the overall strength of the streaming flow, increase as the tip angle decrease [108, 109]. In comparison, microfluidic chamber without any micro-fabricated structure had significantly less cavitation under the same acoustic energy input. In that case, as shown in Fig. 4.2a, cavitation only present at certain area of the edge of the chamber (also see SI video1). This finding is consistent with the aforementioned notion that cavitation was initiated at the interface between hydrophobic surface and the liquid. As we increase the acoustic energy, an interesting behavior of cavitation was observed inside the microfluidic chamber with multiple “crescent” PDMS structures (SI video 2). The cavitation of air bubble travel to a common center point due to complex interplay of primary and secondary Bjerknes forces (Fig. 4.2c)[110]. A simple visual analysis suggests this kind of interplay between different “crescent” structures can further increase and distribute cavitation inside the entire micro-chamber (Fig. 4.2c). We also tested the performance of different waveforms (i.e., square wave, sine wave and triangle wave), generated by the function generator, in terms of the cavitation intensity (SI video 3~5). As shown in Fig. 4.2d~2f, at the same frequency and voltage, the square wave (Fig. 4.2d) generated the highest acoustic cavitation around the “crescent” structure. The sine wave generated slightly lower cavitation around the micro-fabricated crescent structure (Fig. 4.2e). The triangle wave generated the least intensive acoustic cavitation around the structure (Fig. 4.2f). The different acoustic intensity generated by different waveforms is in correspondent with their different root mean square (rms) voltage, as the rms voltage of sine wave and triangle wave is $1/\sqrt{2}$ and $1/\sqrt{3}$ of the rms voltage of square wave.
During sonication, a large amount of heat was generated. The protein structure on the chromatin could be potentially damaged due to overheating[97]. Here we adopted two strategies to do the cooling. First, we placed the whole device on an ice bag during the operation. Secondly, we use a programmed relay to break long continuous sonication duration into several short periods with rest time in between. In order to monitor the sample temperature, we attached a thermal couple on the glass slide, just beneath the microchamber. Since the thickness of glass slide is only 1mm, the real temperature inside the micro-chamber can be truly reflected.

**Fig. 4.1** Schematic of the integrated system incorporating both on-chip sonication and IP. The connection between different devices (e.g. function generation, amplifier, transducer, etc.) was represented using dotted magenta line. The transducer was
bound with microfluidic chip using epoxy glue. The micro-fabricated “crescent” structures, which was able to entrap air inside the micro-chamber, was shown in the Fig..

![Cavitation inside the microfluidic chamber](image)

**Fig. 4.2 Cavitation inside the microfluidic chamber.** (a) Cavitation arising at the edge of the micro-chamber without any micro-fabricated structure. (b) Cavitation initiation in the chamber with multiple “crescent” structures. (c) Intensive cavitation covering majority part of the chamber with “crescent” structures. (d) ~ (f) Cavitation around a “crescent” structure under different waveform: square wave (d), sine wave (e), triangle wave (f). Snapshots of cavitation were obtained using a camera operating at 20fps. An AC voltage of 61 kHz was applied in (a) and (b). An AC voltage of 25V and 61 kHz was applied in (c). An AC voltage of 20 V and 61 kHz was applied in (d)~(f). Scale bar is 0.5mm.

4.3.2 On-chip sonication of crosslinked cells and genomic DNA
Microscale sonicator was previously reported for DNA extraction and fragmentation[111-115]. There is no report about miniaturized sonication of crosslinked cells. Cells was crosslinked using formaldehyde to stabilize the protein-DNA interactions for a variety of important assays, e.g. Chromatin immunoprecipitation (ChIP) and chromatin conformation capture (3C). However, chemical crosslinking makes cells and their nucleus robust and hard to break by enzymatic digestion. Therefore, crosslinked cells are normally fragmented via sonication inside ionic detergent (i.e. 1% SDS) contained sonication buffer, which promotes sonication efficiency considerably [97, 116]. In the following ChIP experiment, reaction buffers containing diluted SDS (~0.1%) are employed, since the high concentration of SDS in the sonication buffer is harmful for the immunoprecipitation assay [97].

As dilution is not an option for our device, here we demonstrate the sonication performance of our protocol with cross-linked mammalian cells using 0.1% SDS contained IP buffer instead of regular sonication buffer with 1% SDS. We cross-linked the cells and introduced them into the microfluidic sonicator for ultrasonic shearing. We took microscopic photos before and after the on-chip sonication, as shown in Fig. 4.3a, b and c. We noticed that, before sonication(Fig. 4.3a), most of the cell maintained their integrity, even though the ionic and non-ionic detergent (i.e. SDS and Triton-X 100) could permeabilize the cross-linked cells. After on-chip sonication for only 1 min(Fig. 4.3b), we could hardly observe any integrated cell except some bright spots in the view, which we assumed as the cell nucleus or cell debris. After on-chip sonication for 2 min, we could not observe any cell nucleus, which suggests complete chromatin releasing from the cell nucleus. The released chromatin can be further sheared into small pieces, as resolved
using gel electrophoresis (Fig. 4.3D~F). During the experiments, we found two important parameters, the sonication intensity as well as the sonication duration, can greatly impact the final fragment size of sheared chromatin. By adjusting these parameters, we can obtain chromatin with wide range of fragment size ranging from several thousand bp to as small as ~100bp. For example, with fixed sonication duration (5min), the fragment size could be significantly reduced using increased sonication intensity (Fig. 4.3d). Similarly, with fixed sonication intensity, better fragmentation of chromatin could be achieved using increased sonication time, as shown in Fig. 4.3E. It should be noted that the optimum fragment size for chromatin immunoprecipitation is around 100bp~700bp.

The video (SI video1) and Fig. data (Fig. 4.2) indicate that micro-fabricated “crescent” structures inside the microchamber can significantly promote the cavitation inside the chamber. The gel electrophoresis result (Fig. 4.3F) further confirmed the correlation between enhanced cavitation and improved chromatin fragmentation. After 5min sonication using 20V, chromatin can be completely sonicated into less than 600bp inside chamber with “crescent” structures. In comparison, sonication inside chamber of no micro-fabricated structures generated a broad smear of chromatin fragments, with a large portion of chromatin remain over 1000bp, under the same sonication condition.

Using our device and protocol, we also sonicated genomic DNA into short fragments. Genomic DNA was prepared off-chip using commercial kit and then loaded into the micro-chamber. Comparing with sonication of crosslinked cell, genomic DNA required less acoustic energy to generate similar fragment size. For example, 2 min sonication with a voltage of 20v is able to completely fragment the genomic DNA into less than 500bp, as shown in the first lane of Fig. 4.3G. Using the same experimental condition, we got DNA
fragments with same fragment size (Fig. 4.3G lane 2), which confirmed the repeatability of our method. Similarly as the sonication of cross-linked cells, smaller DNA fragments can be achieved with increased sonication intensity or duration (Fig. 4.3G lane 4~6). Genomic DNA without sonication was also shown in the Fig. (Fig. 4.3G lane 7) as a control sample.

**Fig. 4.3** The performance of on-chip sonication. (A)~(C) Microscopic picture of crosslinked GM12878 cells before and after on-chip sonication. (A) Before the sonication. (B) Sonication for 1 min. (C) Sonication for 2 min. (D)~(G) Gel electrophoresis results of on-chip sonication. (D) Fragmented chromatin with increased sonication intensity. The voltage applied was 15V, 20V and 25V for lane 1 to lane 3. The sonication duration was 5 min for the all the samples. (E) Fragmented chromatin with increased sonication duration. The sonication duration was 2 min, 4 min, 6 min and 8 min for lane 1~4. The voltage applied was 20V for all the samples. (F) Chromatin sonicated in micro-chamber
with (lane 1) and without (lane 2) “crescent” structures. The same sonication conditions (sonication duration: 5 min, sonication voltage: 20 V) were applied in both cases. (G) Fragmented DNA under different sonication condition: (1) 20 V, 3 min, (2) 20 V, 3 min (3) 15 V, 2 min (4) 25 V, 2 min (5) 15 V, 2 min (6) 15 V, 4 min (7) no sonication. The brightest band in each DNA lander indicates 1 kbp. During sonication, in every 10 s period, there will be 3 seconds on and 7 seconds off. For the sonication of chromatin, 30 K cross-linked cells was used for each condition. For the sonication of genomic DNA, 50 ng genomic DNA was resolved for each condition.

4.3.3 Integrated on-chip sonication followed by ChIP and MeDIP assay

The sonicated chromatin or DNA can be applied for the ChIP or MeDIP assay directly according to steps shown in Fig. 4.4. Briefly, the crosslinked cells or genomic DNA were loaded into the microfluidic chamber for the on-chip sonication (Fig. 4.4A and B). After the generation of fragmented chromatin or DNA, antibody coated beads were loaded into the microfluidic chamber (Fig. 4.4C). The low-intensity acoustic energy was applied to generate acoustic streaming (SI video 6 and 7) without causing any cavitation inside the chamber to enhance the mixing of beads and solution (Fig. 4.4D). During this step, immunoprecipitation (either ChIP or MeDIP) was achieved. After immunoprecipitation, magnetic beads was retained inside the chamber, while IP buffer contained non-specific bound chromatin/DNA or other impurities was flushed out using fresh washing buffer (Fig. 4.4E). The wash process was then initiated with assistance of acoustic streaming (Fig. 4.4F). After a series of washing (Fig. 4.4G), the magnetic beads were finally collected outside the chamber (Fig. 4.4F) for the downstream purification and analysis. The entire
on-chip processing time is around 40 min, of which the IP took only 30 min comparing with overnight incubation by conventional ChIP/MeDIP protocol.

**Fig. 4.** The procedure involved in an integrated microfluidic sonication and IP.

(a) Crosslinked cells were loaded into the chamber. (b) On-chip sonication (c) Antibody coated magnetic beads were loaded into the chamber (d) Acoustic enhanced on-chip
The on-chip sonication generated a large amount of heat, which may potentially damage the protein structure of the chromatin and impact the quality of ChIP DNA. The quality of ChIP DNA was indicated using “percentage input”, which was calculated using qPCR signal (see equation 1) at known positive and negative loci for H3K4me3 (with the primer sequences listed in Supplementary Table 1). The positive locus is where the target epigenetic mark located, while the negative locus has no target epigenetic mark. ChIP DNA with good quality is supposed to have high percentage input at positive loci and low percentage input value at negative loci, which suggest high enrichment of target chromatin. We found that the quality of ChIP DNA were affected by our sonication protocol. We optimized this protocol by adjusting the ratio between on and off time in a fixed time cycle. We obtained the temperature profile (Fig. 4.5a) of the microfluidic chamber, under different sonication condition, as well as the corresponding qPCR data to understand this process. The temperature remained around 4°C, if we sonicated for only 1 second in a totally 10 seconds period (the magenta curve in the temperature profile Fig.). The corresponding qPCR result in Fig. 4.5d indicated high “percentage input” obtained at both two positive loci as well two negative loci. This result suggests both target chromatin and non-target chromatin was immunoprecipitated as a result of incomplete chromatin fragmentation. By increasing the sonication time to 3 seconds in a 10-second cycle, the working temperature inside microfluidic chamber increase to as high as 20°C (blue curve in Fig. 4.5a). The enrichment of ChIP DNA was significantly improved as qPCR signal at
two positive loci (UNKL and C9orf3) were above 25%, while qPCR signal at two negative loci (N1 and N2) were around 1%. Therefore a relative fold enrichment (comparing percentage input value at positive loci against negative loci) of above 25 is achieved in this case. By further increasing the sonication “on” time to 5s in a totally 10 s, the dramatically increased heat inside the micro-chamber drove the temperature to as high as 43°C. With increased sonication, the chromatin is supposed to be better fragmented. The qPCR result after ChIP assay, however, indicated decreased enrichment comparing with previous condition (3s-on and 7s-off). The qPCR signal at positive locus “UNKL” remained high, however, qPCR signal at the other positive locus”C9orf3” dramatically decreased from previous ~25% to less than 10%. Additionally, qPCR signal at two negative loci increased almost 4 fold for N1 (now 4%) and two fold for N2 (now~2.5%). The decreased enrichment or quality of ChIP DNA is associated with elevated working temperature, as 43°C is supposed to be high enough to degrade some protein [117]. If we further increase the sonication time to 9s in a total 10s, the qPCR signal was almost gone at all four loci, either positive or negative. This is easy to explain, with this protocol, the heat accumulated in a very short time without enough time for the heat dissipation via ice bag. Therefore, the temperature easily jumped to as high as 70°C, which is supposed to denature almost all the protein and de-crosslink the bound between protein and DNA [97]. Of all four sonication protocols, the optimal condition is 3s on and 7s off in a 10s cycle. With this protocol, the chromatin was fully fragmented into the working range (100bp~700bp) for ChIP assay, while the generated heat during 3s sonication was mostly dissipated during the 7s off time.
Fig. 4. 5 The correlation between sonication temperature and ChIP result. (a) The temperature profiles of microfluidic chamber under different sonication condition. (b)~(e) : ChIP-qPCR results corresponding to the temperature profile in (a) . An AC voltage of 25V, 60 kHz was applied for 5min during sonication. Cross-linked GM12878 cells were used in these experiments.

With optimized sonication protocol, we sonicated crosslinked cells with different sample sizes ranging from 10k cells to 100 cells. As a proof of concept, we performed on-chip IP (targeting H3k4me3) followed by qPCR analysis, as shown in Fig. 4.6 a. We examined the relative fold enrichment at known positive and negative loci for H3k4me3. For all three cell sample sizes, the fold enrichment at positive loci is significantly higher than that at the negative loci, which suggests high quality of ChIP DNA. In addition, all three data sets have similar enrichment pattern at UNKL and c9orf3. These results
demonstrate that our assay effectively distinguished different levels of histone methylation among different genes.

We also sonicated different amount of genomic DNA (i.e. 50ng, 5ng and 500pg) followed by on-chip MeDIP assay. Primers (with the primer sequences listed in Supplementary Table 2) targeting two positive controls (SNRPN and MAGEA1 promoters) and two negative controls (GABRB3 and GAPDH promoter) were employed for the qPCR data analysis [118]. The relative fold enrichment values at two positive loci were significantly higher than that at the two negative loci. In addition, the enrichment patterns of three data sets at different positive loci are similar, which further confirmed the success of our protocol.

Consistent with that of other publications on ChIP with low cell numbers, the decreased enrichment was noticed using decreased cell numbers or amount of gDNA. This is largely due to dramatically reduced amount of target molecular, which increases the chances of non-specific binding [61]. Nonetheless, the overall enrichment profile was not affected by increased background, even we went down to 100 cells.

![Graphs showing relative fold enrichment](image-url)
Fig. 4.6 qPCR results of ChIP and MeDIP after on-chip sonication. (a) qPCR analysis of ChIP DNA at two known positive loci and two negative loci, using crosslinked cells from 10k cells to 100 cells. An AC voltage of 25V, 60 kHz was applied for 5 min during sonication. Cross-linked GM12878 cells were used in these experiments. (b) qPCR analysis of MeDIP DNA at two know positive control and two negative control, using gDNA ranging from 50ng to 500pg. An AC voltage of 20V, 60 kHz was applied for 3 min during sonication. Cross-linked GM12878 cells were used in these experiments.

Several features of our device may contribute to the high sensitivity and high speed. First, the sonication of cells and ChIP/MeDIP was performed in the same chamber with minimal material loss. Secondly, the working temperature inside the chamber was monitored. An optimal sonication protocol balanced both sonication efficiency and heat dissipation (cooling). Thus, on the one hand, the short chromatin fragments can be obtained, on the other hand, the protein epitopes were preserved. Thirdly, the reaction chamber for immunoprecipitation is small (~7 µl). Such a tiny volume (compared to 1 ml tube used in conventional assays) ensures that the concentration of chromatin was comparable to that in conventional protocols even when the number of cells is $10^4$–$10^5$ times fewer. Additionally, the concentration of antibody coated beads in the tiny chamber was ~100 times higher than that in conventional assays. The close proximity among beads greatly increases the chance for chromatin adsorption on the bead surface due to short diffusion lengths involved and thus speed up the reaction. Finally, we employ acoustic streaming to mix the sample in the steps of both immunoprecipitation and washing. Acoustic enhanced mixing greatly increases the reaction efficiency by (1) promoting rapid transport of target molecular in solution to the diffusion boundary layer.
on the beads surface and thus allowing for continuous replenishment of fresh target molecular around the beads that have been depleted of complementary targets; (2) reducing the thickness of the diffusion boundary layer on the beads’ surface, thus resulting in accelerated antibody-antigen kinetics[119]. Low-amplitude acoustic wave generated inside the chamber also benefit the removal of non-specific binding without compromising the specific binding during the washing steps [120, 121]. We assumed the increased shear stress due to acoustic streaming inside the chamber can efficiently remove majority part of the non-specific binding without removal the specific binding, which has much high affinity and binding strength to the antibody. Moreover, the acoustic based reaction/ washing greatly promotes the degree of automation by eliminating tedious manual handing(e.g. pipetting) involved in conventional ChIP/MeDIP protocols. The automated on-chip procedures were finished within 1 h, and the entire ChIP assay was conducted within 3 h including off-chip protein digestion/DNA purification (0.5 h) and real-time PCR (1.5 h). In particular, the immunoprecipitation step was shortened to 30 min (compared to overnight in traditional ChIP protocols).
CHAPTER 5 Detecting intracellular translocation of native proteins quantitatively at the single cell level

5.1 Introduction

Within eukaryotic cells, proteins efficiently and selectively transit between functionally distinct subcellular compartments including plasma membrane, cytosol, nucleus and other membrane-enclosed organelles. The subcellular localization of an intracellular protein or the change of it (i.e. intracellular translocation) is highly significant for several reasons. First, intracellular translocation can be a prerequisite for proteins to carry out their intended functions. For example, a transcription factor needs to move from the cytosol into the nucleus in order to regulate gene transcription and such events occur typically as a consequence of outside stimuli to the cell. Second, since translocation is often associated with modification and activation at the molecular level (e.g. phosphorylation and methylation), the subcellular localization of the protein molecule is often indicative of its state. In most cases, the proteins are only active at their intended subcellular location. Finally, subcellular mislocalization of proteins leads to diseases ranging from metabolic disorders to cancers[122, 123]. Mislocalization of Akt, NF-κB, FOXO, p27, and p53 have been well-documented as key features in a variety of cancers[123, 124]. Modulation of protein translocation is practiced as an important therapeutic approach for cancer treatment[122, 123]. The subcellular location of a target protein can also serve as a useful read-out for high-content screening of cancer drugs [123].
Conventionally intracellular translocations such as nucleocytoplasmic transport (i.e. the translocation between the nucleus and the cytosol[125, 126]) have been evaluated using fluorescence microscopy or subcellular fractionation[127-134]. However, there are important limitations with these approaches. Fluorescence microscopy (including total internal reflection fluorescence microscopy, or TIRFM[134]) typically analyzes a limited number of cells and does not provide information on the distribution of the cell population. Subcellular fractionation involves lysis and homogenization of cells and then separation of the materials from various subcellular compartments by centrifugation. The data obtained by subcellular fractionation reflect only the average properties of the cell population without revealing the heterogeneity that is often critically involved in cell signaling networks[135-140]. For example, when cells show an all-or-none response to a particular stimulus (bistability), only a subset of cells respond to the signal[136, 141-144]. Thus high-throughput methods are desired to generate information on the translocation of a large number of cells with single cell resolution.

Laser scanning cytometry (LSC) [145, 146] and imaging flow cytometry [147] which permit rapid acquisition of fluorescence images of solid phase or flowing cells, have been used to detect the subcellular localization of intracellular proteins. However, these instruments are very expensive and typically have only limited accessibility through large central facilities. More importantly, these imaging-based instruments require sufficient exposure time in order to generate enough spatial resolution for the analysis. This determines that their throughputs are no more than several hundred cells per second (compared to $10^4$-$10^5$ cells per second by a conventional flow cytometer which detects...
only fluorescence intensity) and the complex algorithms used for image analysis often introduce errors and bias[148-150]. In our previous work, we developed electroporative flow cytometry to examine protein translocations by adding electroporation-based protein release and flow cytometric screening[151, 152]. Unfortunately, electroporative flow cytometry requires special apparatus for combing electroporation and laser-induced fluorescence detection. More importantly, the approach requires proteins of interest tagged by fluorescent protein markers and does not allow examination of native proteins and primary cells isolated from animals and patients. We also explored using TIRF-based flow cytometry for probing protein translocation[153]. However, its potential for high resolution recognition of nucleocytoplasmic translocation was limited.

In this work, we demonstrate a simple method that combines selective chemical release of cytosolic proteins and standard fluorescence immunostaining for detecting the translocation of native proteins at the native expression level with single cell resolution. We demonstrate the proof-of-principle by detecting nucleocytoplasmic transport of an important transcription factor NF-κB. NF-κB undergoes nucleocytoplasmic transport from the cytosol to the nucleus in order to regulate transcription and gene expression, upon extracellular stimuli (e.g. by TNFα)[154, 155]. Briefly, we used saponin (i.e. a class of amphipathic glycosides) to selectively release cytosolic fraction of intracellular proteins by dissolving the cholesterol content and permeabilizing the plasma membrane[156]. Such treatment was followed by fluorescence immunostaining of residual NF-κB. The cell population was then screened by a common flow cytometer for fluorescence intensity of each cell. We showed that the fluorescence intensity of a cell could be correlated to the subcellular localization of the protein. Taking advantage of common flow cytometry which
is widely accessible, our approach detects the translocation of native proteins without imaging and with high throughput. Our approach is readily compatible with analysis of primary samples from animals and patients.

5.2 Materials and methods

5.2.1 Cell sample preparation

HeLa (CCL-2) cells were grown at 37 °C with 5% CO₂ in Dulbecco’s modified Eagle’s medium (DMEM) (Mediatech, Herndon, VA) supplemented with 10% (v/v) fetal bovine serum (Sigma) and 1% penicillin (100 mg/ml, Sigma). The cells were trypsinized and diluted at a ratio of 1: 5–1: 8 every 2 d to maintain the cells in the exponential growth phase. The harvested cells were centrifuged at 300 g for 5 min and resuspended in DMEM culture medium at a final concentration of 1×10⁶ cells/ml before experiments. To stimulate cells, cells (1×10⁶/ml) were suspended in the culture medium with various concentrations of TNF-α (AbD Serotec, Raleigh, NC, USA) at 37°C for various periods.

5.2.2 Standard fluorescence immunostaining

Fluorescence immunostaining was conducted following the literature with minor changes [157]. The cell sample (~1×10⁶ cells) was fixed in 100 μl pre-warmed (at 37°C) fixation buffer (4% paraformaldehyde in PBS buffer) for 10 min. Subsequently, the fixed cells were washed with a blocking buffer (1% BSA in PBS buffer) and permeabilized with 100 μl of a permeabilization buffer (0.2% Triton X100 in PBS buffer). After 20 min incubation with the permeabilization buffer, the cells were centrifuged at 300 g for 5 min to remove the permeabilization buffer and washed with the blocking buffer. The cells were
then incubated with the blocking buffer containing a primary antibody [1:100 dilution of NF-κB p65 (sc-8008, Santa Cruz Biotechnology, Dallas, Texas, USA)] for 1 h at room temperature. After incubation with the primary antibody, the cell sample was pelleted by centrifugation (300g, 5 min) and washed twice with the blocking buffer. Then, the cells were incubated (protected from light) in the blocking buffer containing fluorophore-conjugated secondary antibody [1:150 dilution of DyLight™ 488 Goat anti-mouse IgG1 antibody (409102, Biolegend, San Diego, CA)], which bound to the primary antibody, for 1 h at room temperature. The staining solution was then aspirated out and the labeled cells were washed twice with the blocking buffer to remove nonspecific binding. The cells were stored in PBS with 0.1% sodium azide at 4°C, if not immediately analyzed by flow cytometry.

5.2.3 Selective release of cytosolic proteins followed by fluorescence immunostaining

The cell sample (~1×10^6 cells) was incubated in 100μl of the releasing buffer [0.05% (w/v) saponin (ID# 419-25A, Chem Service, West Chester, PA, USA) in DMEM] for 5 min. The processed cells were then pelleted by centrifugation at 300g for 5 min to remove excessive releasing buffer and then immediately fixed by the fixation buffer for 30 min. Subsequently, the fixed cells were permeabilized in 100 μl of the permeabilization buffer. After 5 min incubation in the permeabilization buffer, the cells were centrifuged at 300 g for 5 min to remove excessive permeabilization buffer and washed once with the blocking buffer. The rest of the procedure involving labeling using primary and secondary antibodies was the same as that in “standard fluorescence immunostaining”.
5.2.3 Fluorescence microscopy

Immunostained cells were transferred to a 96 well plate and then centrifuged for 5 min at 300×g to settle the cells to the bottom. Fluorescence images were taken by an inverted fluorescence microscope (IX-71, Olympus, Melville, NY) with a 20X dry objective (0.5 NA). The fluorescence excitation was provided by a 100W mercury lamp. The excitation and emission were filtered by a fluorescence filter cube (exciter HQ480/40, emitter HQ535/50, and beam splitter Q505lp, Chroma Technology) for observing green fluorescence.

5.2.4 Western blotting

Cellular protein samples were made using NE-PER Nuclear and Cytoplasmic Extraction Reagents (Pierce Biotechnology, Rockford, IL, USA) following the manufacturer’s recommendations. The total protein concentration for each sample was measured by Pierce BCA protein assay kit (Pierce Biotechnology, Rockford, IL, USA), and then equal amount of denatured proteins from each sample was separated by standard SDS-PAGE and analyzed by Western blotting. Briefly, samples were loaded into polyacrylamide gel for SDS-PAGE followed by transferring to a polyvinylidene fluoride (PVDF) membrane, which was blocked by 5% milk in TBST (Tris-Buffered Saline, 0.1% Tween-20) buffer. The membrane was stained with 1:1000 diluted primary antibody (SC-8008, Santa Cruz Biotechnology) and secondary antibody [Rabbit anti-mouse horseradish peroxidase (HRP) (Pierce, Rockford, IL)] for 1h each. Membranes were visualized by LAS-3000 luminescent image analyzer (Fujifilm, Hanover Park, IL, USA) after chemiluminescence treatment with Pierce ECL Western Blotting Substrate (Pierce
Biotechnology, Rockford, IL, USA). The intensity of each band was quantified using ImageJ software.

5.2.5 Flow cytometry analysis

Fluorescently stained cell samples were analyzed at medium flow rate by a FACS Canto II cytometer (BD, San Jose, CA, USA) equipped with 488 nm laser/filters for FITC and forward scatter (FSC) detection. For each histogram, 10000~20000 events were collected. The cytometer was routinely calibrated with Calibrite beads (BD). Stained cell samples may be stored for up to 24 h at 4°C before analysis without a significant loss in fluorescence intensity. The data were processed by FlowJo and Origin 9.0.

5.3 Results and Discussion

In Fig. 5.1, we outline the procedure and working principle of our protocol (i.e. selective release and immunostaining), in comparison to those of the standard immunostaining protocol, and how these procedures generate different results for single cell screening with standard flow cytometry.
**Fig. 5.1** The comparison between flow cytometric screening of cells after (A) Standard fluorescence immunostaining and (B) Selective release of cytosolic proteins followed by immunostaining. There is no difference in the fluorescence intensity between cells with and without nucleocytoplasmic transport with standard fluorescence immunostaining, whereas the fluorescence intensity of single cells is correlated with the protein subcellular localization (or the activation state) with our selective release combined with immunostaining protocol.

In our experiments, we applied TNFα to stimulate cells and produce NF-κB translocation from the cytosol to the nucleus. With standard fluorescence immunostaining, HeLa cells with NF-κB translocation (i.e. +TNFα) and those without NF-κB translocation (i.e. untreated) are crosslinked by paraformaldehyde and permeabilized by Triton X 100 before all intracellular NF-κB is fluorescently labeled by antibodies that specifically target
NF-κB. The crosslinking by paraformaldehyde conserves all proteins in the cells and the permeabilization ensures the full access of the targeted protein (i.e. NF-κB in this study) by the antibodies for labeling. When these cells are subsequently screened by flow cytometry, which detects fluorescence from the entire cell, the translocation does not create difference in the fluorescence intensity detected because it only varies the subcellular localization of the protein, not the overall expression level (Fig. 5.1A). In comparison, in our method, we add a selective release step before immunostaining. During this step, saponin is used to treat the cells to dissolve cholesterol in the plasma membrane and make the membrane leaky[156, 158, 159]. Saponin is known to primarily permeabilize the plasma membrane while keeping the cholesterol-poor internal membranes (e.g. the mitochondrial membrane and nuclear envelope) intact[156, 160]. Gentle treatment by saponin has also been shown to have minimal effects on cellular functions such as protein synthesis[161, 162]. Cytosolic proteins are preferentially released out of the cells and the nucleic proteins are largely unaffected in their amounts. Thus cells with NF-κB translocation (with NF-κB mostly in the nucleus) have far less decrease in the NF-κB amount due to the release, compared to the cells without NF-κB translocation (with NF-κB primarily in the cytosol). The selective release step is immediately followed by immunostaining that fixes the cells and labels all the remaining intracellular NF-κB. The flow cytometry results now are significantly different for cells with translocation and those without translocation, with the latter showing much smaller fluorescence intensity. Thus we are able to link the subcellular localization of the protein with the detected fluorescence intensity of a cell treated by our protocol.
Fig. 5.2 shows the flow cytometry data obtained after standard immunostaining (Fig. 5.2A and 4.2B) and combined selective release and immunostaining (Fig. 5.2C and 4.2D).

**Fig. 5.2** The detection of NF-κB translocation using the selective release/immunostaining protocol and conventional flow cytometry. The cell populations with and without TNFα stimulation were examined. The difference between the two populations was not revealed in either the fluorescence intensity histograms (A) or 2D dot plots (B) involving both fluorescence intensity and forward scatter (i.e. FSC), when standard immunostaining was used. In comparison, there was pronounced difference between the two populations in both the fluorescence intensity histograms (C) and 2D dot plots (D) when selective release followed by immunostaining was conducted. TNFα stimulation was conducted at 37°C with 50ng/ml TNFα for 30min. Selective release was performed using 0.05% saponin for 10 min at room temperature.
Fig. 5.2A shows that as expected, with standard immunostaining the fluorescence intensity histogram generated by a cell population without NF-κB translocation overlaps with that generated by a cell population with the translocation (stimulated by TNFα for 30 min). In comparison, the two cell populations exhibit marked difference in the fluorescence intensity histogram when we had selective release of cytosolic proteins by saponin (0.05% saponin for 10 min) before the immunostaining (Fig. 5.2c). Furthermore, we also discovered that the two cell populations (processed with the selective release protocol) were even more distinct when two-dimensional dot plots were used to include information on the cell size (via detecting the forward scatter signal) (Fig. 5.2D). The two cell populations were entirely separated from each another in Fig. 5.2D. This improvement is attributed to the differentiation of large cells without translocation (NF-κB in the cytosol) and small cells with translocation (NF-κB in the nucleus). These two subpopulations may have similar fluorescence intensities after the selective release protocol but are very different in the cell size. The fluorescence images of cells after standard immunostaining and selective release/immunostaining were also collected (Fig. 5.S1). The images confirm the proposed mechanism in Fig. 5.1. Our technique renders the fluorescence intensity different for cells with translocation and those without translocation, whereas standard immunostaining reveals the different localizations of the protein via imaging with the overall fluorescence intensity from whole cells being the same for the two types of cells.
**Fig. 5.S 1 Fluorescence images of untreated cells** (A, C) and TNFα-stimulated cells (B, D) after standard immunostaining (A, B) and our selective-release-based immunostaining (C, D). TNFα stimulation was performed by adding 50 ng/ml TNFα for 30 min at 37°C. Selective release was performed by 0.05% saponin for 10 min at room temperature. Scale bar =100μm.

We also optimized the selective release protocol in order to create the maximum differentiation of cells based on NF-κB subcellular localization. We varied the concentration and duration of saponin treatment for two cell populations (untreated cells and cells stimulated by 50ng/ml TNFα for 30min) and observed the difference in their fluorescence intensity histograms (Fig. 5.S2). Fig. 5.S2A shows that with 10 min treatment time, the saponin concentration of 0.05% yielded the best separation between the two cell populations. The decreased differentiation at higher saponin concentrations (0.2-0.5%) presumably resulted from the release of both cytosolic and nucleic proteins. Fig.
5.S2B shows that with a fixed concentration of 0.05% for saponin, the optimal treatment time was in between 1 and 10 min. Exceedingly long treatment times (>10 min) also led to decreased separation, due to release of both cytosolic and nucleic proteins. Thus we determine that saponin treatment of 0.05% concentration and 1-10 min works the best as the selective release step for differentiation of cells with NF-κB in the nucleus and those with the same protein in the cytosol.

![Graph showing optimization of conditions for maximal differentiation of untreated and TNFα stimulated cells using selective-release-based immunostaining and flow cytometry.](image)

**Fig. 5.S2** The optimization of conditions (saponin concentration and treatment duration) for maximal differentiation of untreated and TNFα stimulated cells using selective-release-based immunostaining and flow cytometry. **TNFα stimulation** was conducted by incubating cell with 50 ng/ml TNFα at 37°C for 30 min. Fluorescence histograms were obtained after selective-release-based immunostaining and flow cytometric screening. (A) Various saponin concentrations used when the treatment duration was 10 min. (B) Various treatment durations used while the saponin concentration was 0.05%.
Using the optimized selective release protocol (in combination with immunostaining), we show that our method is effective for revealing various degrees of NF-κB translocation in the cell population. In Fig. 5.3, various concentrations of TNFα were used to stimulate the cell population for 30 min. With low concentrations (0.1-1.0 ng/ml) of TNFα, the cell populations appear to have lower degree of NF-κB translocation overall and a subpopulation of cells have no translocation at all (based on the broad peak shape which is not Gaussian).
**Fig. 5.3** The dose dependence of TNFα stimulation analyzed by our approach. TNFα stimulation was conducted at 37°C for 30 min. Selective release was performed by incubating cells with 0.05% saponin for 10 min at room temperature.

With high concentrations of TNFα (10-100 ng/ml), the translocation occurs more completely for the cell population and the separation between the stimulated population and the control is increasingly complete.

Finally, we used our approach to examine temporal dynamics in the cell population during NF-κB translocation. In Fig. 5.4A, the shift in the fluorescence intensity histogram suggests the movement of NF-κB from the cytosol to the nucleus after TNFα stimulation. The increase in the fluorescence intensity of the histogram indicates the increased occupation of the nucleic localization over time. The data show that the translocation occurs substantially within the first 5 min after simulation. There is increased translocation until 30 min after stimulation. Interestingly, translocation in the reverse direction (from the nucleus to the cytosol) occurs between 40-60 min. Such reverse translocation was previously reported in the literature [163] and is due to re-inhibition of newly synthesized repressor IκB [129]. The western blotting analysis (Fig. 5.4B) also corroborates these findings by our technique.
Fig. 5.4 The temporal dynamics of NF-κB translocation detected by our approach (A) and verified by western blotting analysis of the nuclear fraction (B). Cells were stimulated by 50 ng/ml TNFα at 37°C for various durations (0 ~ 60 min). Selective release was performed by 0.05% saponin for 10 min at room temperature.
The resolution of the technology depends on several factors. First, the amount of the protein translocation between the nucleus and the cytosol affects the resolution. Based on the comparison of Fig. 5.4a and Fig. 5.4b, the translocation of ~13% of the total NF-κB in the entire cell can be clearly resolved by the flow cytometry data. Second, the completeness and selectiveness of cytosolic release by saponin are critical for high resolution. As demonstrated in Fig. 5.S2, optimal treatment conditions need to be obtained for a particular cell/protein system in order to reach the best resolution. Third, the immunostaining after the selective release needs to be complete and yields strong fluorescence signal. This facilitates obtaining high quality flow cytometry data.

To conclude, by combining selective release of cytosolic proteins via chemical permeabilization with fluorescence immunostaining, we develop a protocol that links the fluorescence intensity of a single cell with the subcellular localization of a targeted protein. By screening the fluorescence emitted by single cells using common flow cytometry, we are able to detect the translocation quantitatively with single cell resolution. Because fluorescence immunostaining is ideally suited for studying cell samples from animals and patients, our approach provides a very simple route for examining protein translocation at the single cell level with direct biomedical relevance. We expect that this approach can be extended to a wide range of cell types and proteins.

5.4 Conclusions

The intracellular localization and movement (i.e. translocation) of proteins are critically correlated with the functions and activation states of these proteins. Simple and accessible detection methods that can rapidly screen a large cell population with single
cell resolution have been seriously lacking. In this report, we demonstrate a simple protocol for detecting translocation of native proteins using a common flow cytometer which detects fluorescence intensity without imaging. We sequentially conducted chemical release of cytosolic proteins and fluorescence immunostaining of a targeted protein. The detected fluorescence intensity of cells was shown to be quantitatively correlated to the cytosolic/nuclear localization of the protein. We used our approach to detect the translocation of native NF-κB (an important transcription factor) at its native expression level and examine the temporal dynamics in the process. The incorporation of fluorescence immunostaining makes our approach compatible with the analysis of cell samples from lab animals and patients. Our method will dramatically lower the technological hurdle for studying subcellular localization of proteins.
CHAPTER 6 DROPLET SORTING BASED ON THE NUMBER/TYPe OF ENCAPSULATED PARTICLES USING A SOLENOID VALVE

6.1 Introduction

Droplet microfluidics focuses on transporting and manipulating monodisperse aqueous droplets within carrier oil stream in microfluidic devices [164-166]. Microfluidic droplets provide uniform microscale compartments that are isolated by the water/oil boundary for conducting reactions [167-171] and sorting particles [172-174]. Droplet microfluidics has been increasingly used for encapsulating and processing microbeads, cells, and molecules for high throughput screening with biological applications [175, 176]. For example, microbeads were encapsulated in droplets to create gene encoded protein libraries [177], genomic libraries [178], and conduct single cell genetic analysis [179]. Cells may also be encapsulated into droplets, which enable ultrasensitive detection of secreted molecules [176, 180], drug screening [173], genetic analysis [181, 182], and enzymatic tests [172].

The generation of droplets with encapsulated particles is an important step for these screening applications. Most of encapsulation works are based on random loading that produces droplets from particle laden flow [174, 176, 183-185]. Such process is dictated by Poisson statistics and yields droplets containing uneven number of particles (e.g. with a large fraction of droplets empty when single occupancy is desired). Controlled encapsulation methods have been explored in recent years. Particle alignment offered the potential to generate uniform occupancy beyond Poisson statistics [186-188]. However, the particle ordering occurs only under specific flow conditions and the encapsulation
result could still be affected by the uniformity of the cell distribution throughout the liquid volume. Passive self-sorting of droplets (of 3.5-10 μm diameter) has been demonstrated to separate cell-encapsulated droplets from empty droplets based on the difference in the size[174]. Such scheme does not apply to larger droplets in which the encapsulated particles do not introduce change in the droplet size. Thus, the majority of the droplet screening systems still need to deal with droplets with inhomogeneous particle occupancy. In contrast, most fluorescence-based droplet detection and sorting systems are based on differentiation of the fluorescence intensity[172, 189, 190]. Droplet sorting methods using the number of encapsulated particles as the criterion are in great need.

In this report, we demonstrate a simple one-layer microfluidic device for droplet sorting based on the number of particles encapsulated. In our device, droplets with encapsulated particles flowed through a narrow interrogation channel where the particles inside were detected by laser-induced fluorescence (LIF) one by one. A programmable microcontroller was used to rapidly process the signal and then make the decision for sorting. We chose to use pulsed pressure generated by an off-chip fast-acting solenoid valve for deflecting selected droplets into the collection outlet. This deflection mechanism avoided complexity in the design and fabrication and potential damage to cell viability associated with other sorting mechanisms such as dielectrophoresis, electrokinetic actuation, surface acoustic waves, piezoelectric force and ultrasound beam [172, 191-200]. We demonstrated a sorting speed up to ~30 droplets/s. We tested both microbeads and cells as the encapsulated particles. We demonstrated highly specific sorting of two-particle droplets from a mixed population and the collection yielded very high percentage of the targeted droplets (96.3% for microbeads and 93.7% for cells). We envision that our
device will find applications mainly in two fronts: 1. Generation of uniform droplet population with the same number of particles encapsulated in each droplet; 2. Analysis of a mixed droplet population that yields information on the particle number in each droplet.

6.2 Materials and methods

6.2.1 Microfluidic chip fabrication

The microfluidic droplet sorter was fabricated using standard soft lithography method as described in our previous publication [201]. The microscale patterns were first created using a computer-aided design software (FreeHand MX) and then printed out on high-resolution (5080 dpi) transparencies. The transparencies were used as photomasks in photolithography on a negative photoresist (SU-8 2075, MicroChem, Newton, MA, USA). The thickness of the photoresist and hence the depth of the channels was 50 µm. The pattern of the microfluidic network in the photomask was replicated in SU-8 after exposure and development. The microfluidic chip was molded by casting a layer (about 5 mm thick) of PDMS prepolymer mixture (General Electric silicone RTV 615, MG chemicals) with a mass ratio of A:B = 10:1 on the photoresist/silicon wafer master. The prepolymer mixture was cured at 85 ºC for 2 h in an oven and then peeled off from the master. Glass slides were cleaned in a basic solution (H2O:NH4OH (27%):H2O2 (30%) = 5:1:1, volumetric ratio) at 75℃ for 1 h and then rinsed with DI water and blown dry. The surfaces of the PDMS chip and a clean glass slide were oxidized using a plasma cleaner (Harrick Plasma, Ithaca, NY, USA). The PDMS chip was then immediately brought into contact against the slide after oxidation to form irreversible bonding. The device was baked at 85 ºC for 1 h for further strengthening of the bonding. To render the channel hydrophobic for droplet
generation and manipulation, Aquapel (PPG Industries, Pittsburgh, PA, USA) was used to coat the channel before it was blown out of the channel by air [185].

6.2.2 Fluorescent beads and cell samples

Fluorescent polystyrene beads with diameter of 16±0.05 μm (s.d.) were purchased from Duke Scientific. The density of beads is 1.05 g/ml according to the manufacturer’s specification. Before experiments, the beads were diluted and sonicated in DI water at a concentration of 1×10^6/ml in the presence of 5% (v/v) Pluronic F127 (Sigma, St. Louis, MO, USA), which acted as a steric stabilizer of the beads [202]. The buffer also contained dextran (Sigma, St. Louis, MO, USA) at a weight/volume fraction of 10% in order to increase the buffer density and prevent beads settling during operation. HeLa (CCL-2) cells were grown at 37°C with 5% CO2 in Dulbecco’s modified Eagle’s medium (DMEM) (Mediatech, Herndon, VA) supplemented with 10% (v/v) fetal bovine serum (Sigma) and 1% penicillin (100 µg/ml, Sigma). The cells were trypsinized and diluted at a ratio of 1:5-1:8 every 2 d to maintain the cells in the exponential growth phase. The harvested cells were centrifuged at 300g for 5 min and resuspended in DMEM culture medium at a final concentration of 1×10^6 cells/ml. The cell suspension was incubated with 2 µM Calcein AM green (Invitrogen, Grand Island, NY, USA) for 30 min at 37°C for fluorescent labeling[203, 204].

6.2.3 System setup and operation

A schematic of the droplet sorting device is shown in Fig. 6.1. The microfluidic chip was mounted on an inverted microscope (IX-71, Olympus, Melville, NY, USA) equipped with a 10x dry objective (NA=0.60). The two inlets of the droplet generation section (one
for oil and the other for aqueous particle suspension) (Fig. 6.1A) were connected to syringe pumps (Fusion 400; Chemyx, Stafford, TX, USA) through perfluoroalkoxyalkane (PFA) tubing (1622L; IDEX Health &Science LLC, Oak Harbor, WA, USA). The inlet of the actuation channel was connected to a pressure source of ~1 psig via a solenoid valve (ASCO Scientific, Florham Park, NJ, USA) and PFA tubing. Droplets were produced in the T-junction channel as previously described [184]. As a result of the shear forces that arose by bringing two immiscible fluids together, the subsequent capillary instability led to generation of monodisperse aqueous droplets in the oil. The continuous phase was of fluorocarbon oil (FC-40, 3M, St. Paul, MN, USA) containing 5.0% (w/w) PFPE-PEG block-copolymer surfactant used in our previous publication[185]. To avoid the hydrophilization of the channel, the continuous phase was first introduced to fill all the channels. Droplet movement could be monitored by a CCD camera (ORCA-285, Hamamatsu, Bridgewater, NJ, USA) or a high speed camera (MotionXtra N3, IDT, Tallahasse, FL, USA) mounted on an eyepiece port of the inverted microscope.

A 488 nm laser beam (Cyan-488-50-FP-CDRH, Newport, Franklin, MA) was focused in the narrow interrogation channel by a 10x objective for laser-induced fluorescence detection (Fig. 6.1B). The intensity of the laser beam was adjusted by a set of neutral density filters (Newport, Franklin, MA). The laser was introduced into the microscope via a laser side port and was reflected by a 505DCLP dichroic beam splitter (Chroma Technology, Bellows Falls, VT, USA) into the 10x objective. The fluorescence emission was collected by the same objective and filtered by a spectral filter (D535/40 emission filter, Chroma Technology). A 20 µm pinhole (P20S, Thorlabs, Newton, NJ, USA) was placed in the optically conjugate plane in front of a photomultiplier tube (R9220,
Hamamatsu, Bridgewater, NJ, USA) biased at 550 V, to eliminate out-of-focus signal. The electrical signal generated by the PMT was amplified and filtered by a low noise current preamplifier (SR570, Standard Research System, Sunnyvale, CA, USA) for optimum Signal-to-Noise Ratio (SNR). The signal was then processed by a differential comparator (TL712, Texas Instrument, Dallas, TX, USA) before input into a microcontroller (AT89S51, Atmel, San Jose, CA, USA). When a target droplet was detected, the microcontroller activated a reed relay (R56-1D.5-6D, NTE Electronics, Bloomfield, NJ, USA), which in turn operated the solenoid valve and thus diverted the target droplet into the collection channel. Both the fluorescence analog signal out of the preamplifier and converted digital signal out of the comparator could be examined or recorded by an oscilloscope (B&K precision2530, Newark, Chicago, IL, USA) or a PCI data acquisition card (PCI- 6254, National Instruments, Austin, TX, USA).

6.3 Results and discussion

The microfluidic device, as shown in Fig. 6.1, has a simple one-layer structure. We used a T-junction (Fig. 6.1A) to produce droplets with encapsulated particles, having the carrier fluorocarbon oil and the aqueous particle (microbead or cell) suspension coming from separate inlets. The generated droplets were directed into a narrow interrogation channel, which was only 23 µm wide in order to ensure the vast majority of encapsulated particles detected individually in the laser focal volume [205] (~15 µm in the diameter, generated by a 488 nm laser focused through a 10x objective) (Fig. 6.1B). Droplets were sorted in the downstream where the interrogation channel met with three other channels (i.e. the actuation channel, the waste channel and the collection channel) (Fig. 6.1C).
Without pulsed pressure from the actuation channel, the difference in the hydrodynamic resistance resulted from the different channel lengths (the waste and collection channels were 2 and 3.4 mm long, respectively) drove droplets exclusively into the waste channel. This directional motion can be explained by Zweifach-Fung effect, according to which the asymmetric distribution of pressure gradient and shear forces on the surface of a droplet draw it towards the branch with higher flow rate[206]. The waste and collection channels had large widths (~300 µm) in order to decrease the impact of the droplets on the hydrodynamic resistance [207-209]. The actuation channel was connected to a pressure source (~1 psig) via an off-chip fast-acting solenoid valve operated by a microcontroller. As the fluorescence emitted by encapsulated particles was collected and processed by the optical detection and signal processing system (Fig. 6.1D), the microcontroller made the decision about sorting to the collection channel in case of detecting target droplets (further detailed below). When actuated, the solenoid valve was rapidly switched open and then shut, creating pulsed pressure (~1 psig) with a duration of 7 ms. As we show in the COMSOL simulation(Fig.5S.1), the resulted fluid pressure would sweep the droplet (exiting the interrogation channel) into the collection branch. Our simulation shows that the entire sorting process would take less than 8 ms until the flow restores to its original state.
Fig. 6S.1 *The simulation of the deflection process by pressure actuation by COMSOL Multiphysics 4.2.* The simulation is simplified by only examining the oil phase. The video shows the process occurring in 15 ms. The pressure pulse (1 psig and 7 ms duration) was applied at the actuation channel at 4-10 ms. The color scale indicates the velocity magnitude. The video shows that the pressure pulse would sweep a droplet at the exit of the interrogation channel (the vertical narrow channel) into the collection channel (the left wide channel).

With a response time of roughly 10 ms for the solenoid valve used (based on the manufacturer specification), our system has the potential of sorting up to 55 Hz. It is worth noting that the fastest solenoid valves these days may have response time as short as 1 ms thus would offer much higher sorting rate. The use of solenoid valves for microfluidic sorting makes the device design drastically simple. Furthermore, such arrangement is
entirely compatible with a large number of integrated microfluidic systems because solenoid valves are widely used for two-layer valve switching [210-212].

Fig. 6.1 Schematic of the droplet sorting system. All the microfluidic channels have a uniform depth of 50 µm. (A) Droplet generation. The aqueous particle suspension (containing either microbeads or cells) meets the fluorocarbon oil stream at the T junction to produce droplets. The channel widths of aqueous and oil inlets are 30 and 40 µm respectively. (B) Droplet interrogation. The narrow width (~23 µm) of the interrogation channel allows encapsulated particles to be detected individually (with the image of a droplet containing a single bead shown in the inset). (C) Droplet sorting. The droplets exclusively flow into the waste channel due to the difference in the hydrodynamic
resistance (The length of the waste channel is 2 mm while that of the collection channel is 3.4 mm; both channels are 300 µm wide) until they are swept into the collection channel by the pressure pulse from the actuation channel. The actuation channel (80 µm wide and 2 mm long) is connected to a pressure source (~1 psig) via an off-chip solenoid valve. In the schematic, we suggest selective isolation of droplets that contain two particles from a mixture. (D) Off-chip detection and control system. The fluorescence signal is detected by a photomultiplier tube and further processed by a comparator and a microcontroller. The microcontroller makes the sorting decision and operates the solenoid valve.

We established the robustness of the solenoid valve sorting by demonstrating simple separation of droplets with and without fluorescence. We first examined a mixed population of fluorescent and nonfluorescent droplets generated in the upstream of the sorter by previously published method [213]. In order to facilitate differentiation under phase contrast microscopy, the fluorescent droplets contained fluorescein and the nonfluorescent droplets were doped with trypan blue dye. As shown in Fig. 6.2a, the sorter was able to deliver the nonfluorescent droplets into the waste channel (left) and the fluorescein-containing droplets into the collection channel (right). In general, the accuracy was very high for such sorting. In Fig. 6.2b, we show that the sorting accuracy remained at 97.5% even we increased the droplet throughput to 30 droplets/s. Furthermore, we demonstrate that droplets with encapsulated fluorescent particles can be separated from droplets without particles. As shown in Fig. 6.2c, the empty droplets were directed into the waste channel and the droplets with fluorescent polystyrene beads (D~16 µm) were sorted into the collection channel. The sorting accuracy in this case, varying from 97% at

124
5 droplets/s to 90% at 30 droplets/s (Fig. 6.2d), was slightly lower than that of sorting fluorescein droplets (Fig. 6.2b). This was due to missed detection of encapsulated particles when the small laser focal volume did not fully cover the space occupied by the particle(s). Thus the mistaken sorting events were all false negatives (i.e. taking droplets with fluorescent beads as those without beads). Finally, we also tested droplets with encapsulated cells (Calcein AM stained HeLa cells), as shown in Fig. 6.2e and 2f. Culture medium instead of DI water was used in the aqueous phase and a block-copolymer surfactant was added to the continuous oil phase to stabilize the droplets while keeping cell viability [185]. The sorting accuracy of 95.1% at 5 droplets/s and 86.9% at 30 droplets/s for cell-containing droplets (Fig. 6.2f) is slightly lower than that of sorting bead-containing droplets (Fig. 6.2d). We attribute the difference to the uneven size and fluorescence intensity from individual cells. Despite of this, the accuracy is higher than most of previously described microfluidic FACS sorters [214, 215]. Similar to the bead-containing droplet sorting, there were only false negatives and a high purity of cell-containing droplets (100%) was obtained in the collection end.
Fig. 6.2 Droplet sorting based on emitted fluorescence. CCD camera images confirm that droplets containing (a) fluorescein; (c) fluorescent beads; (e) fluorescently stained HeLa cells were sorted into the collection channel (right). The scale bar is 100 μm. The sorting accuracy (i.e. the percentage of droplets being correctly sorted) in these three cases: (b) fluorescein; (d) fluorescent beads; (f) fluorescently stained HeLa cells, at various droplet throughputs (5-30 droplets/s) is also shown. The flow rates of the aqueous phase and the oil phase were varied in the ranges of 0.05-0.5 μl/min and 2-10 μl/min, respectively, in the experiment to achieve different droplet throughputs. The droplet
diameter ranged from 30 to 120 μm (viewed in the channel). The experiments were conducted in triplicate. For each run, at least 2000 droplets were observed for calculating of the sorting accuracy.

In order to sort droplets based on the number of encapsulated particles, we first examined whether our system generated reliable signal that would indicate the number of particles in each droplet. As shown in Fig. 6.3a, typical fluorescence signal detected from a mixed droplet population with encapsulated fluorescent beads presents peaks of various intensities. As shown in the zoomed-in spectrum in Fig. 6.3a, each droplet generates a broad peak due to the light scatter at the water/oil interface. The broad peaks may have one or two embedded high-intensity peaks that were produced by fluorescent beads encapsulated. The variance in the peak intensity was likely to be due to the randomness in the bead position when each bead passed through the laser focal volume. Similar data were also generated by cell-containing droplets (Fig. 6.3c). In general, the raw fluorescence peaks generated by encapsulated cells showed even more variance than beads due to the nonuniform cell size. Nevertheless, we could use a custom MATLAB program to extract information on the number of particles in each droplet and then generate the occupancy distributions (Fig. 6.3b and 3d). The occupancy distributions produced by both fluorescent beads and stained cells can be fit by Poisson distribution very well. Because the particle occupancy of droplets is known to be governed by Poisson distribution[172, 174, 176, 183], the agreement between the experimental data and the predicted distribution confirms the system’s ability to accurately detect the vast majority of the encapsulated particles.
Fig. 6.3 The fluorescence signal contains information on the number of encapsulated particles per droplet. The fluorescence intensity data over time are shown for droplets with encapsulated (a) fluorescent beads and (c) fluorescently stained HeLa cells. The flow rates were set at 0.15 μl/min and 3.6 μl/min for the aqueous phase and the oil phase, respectively. Using a MATLAB program, we extracted the particle number per droplet from these fluorescence intensity data. The probability distribution in terms of the particle number per droplet obtained from these experimental data (b) fluorescent beads and (d) fluorescently stained HeLa cells (by analyzing droplet populations >2000) fits Poisson distribution (delineated in red line). Poisson distribution
is defined by 

\[ P(k) = \frac{\lambda^k e^{-\lambda}}{k!} \]

, where \( k \) is the number of particle per droplet (0~3) and \( \lambda \) is the average number of particles per droplet (0.39 for beads, 0.44 for cells).

We constructed sophisticated signal processing system including a comparator and a microcontroller for extracting information on the number of particles in each droplet and making sorting decision accordingly, all in real time. We used a comparator to compare the detected fluorescence signal (in analog voltage) with a threshold voltage and the comparator output a constant voltage over the period that the fluorescence signal was higher than the threshold (the output is zero during other times) (Fig. 6S.2).

![Comparator Circuit](Fig. 6S.2 Diagram of the comparator circuit.)

With the preset threshold (higher than the broad peak intensity), the broad peaks generated by light scatter were filtered out and the high-intensity peaks were converted to peaks of uniform height (Fig. 6.4). The digitalized peaks (Fig. 6.4b) were then input into
a microcontroller. A microcontroller is a small computer on a single integrated circuit that is often used in automatically controlled systems. In our case, the microcontroller did the computation to determine two things: 1. Are the consecutive particles encapsulated in the same droplet? 2. How many particles are in a droplet? Fig. 6.5 shows the algorithm used to sort specifically droplets that contain two particles. The algorithm determines whether multiple particles are within the same droplet based on the interval between the particles. When the particles are in the same droplet, the interval is shorter than the time required for a droplet to pass through the detection point, assuming that the distance between droplets was significantly longer than the droplet length. The algorithm also included a counter to register the number of particles in one droplet.

**Fig. 6.4 The processing of the fluorescence intensity data by the comparator.** The data were generated using droplets with encapsulated fluorescent beads and the flow rates of 0.15 μl/min and 3.6 μl/min for the aqueous and the oil phases, respectively. (a)
The raw fluorescence intensity data over time before the processing. A threshold of 0.35 V is set for determining which peaks are generated by fluorescent beads. (b) The signal after the comparator processing. All peaks higher than the threshold are converted to peaks of uniform intensity (5 V).
Fig. 6. A flowchart that summarizes the logic of the microcontroller for sorting two-particle droplets out of a mixed population. A timer and a counter are needed to determine how many particles are in one droplet. The droplet transit time (i.e. the time required for a droplet to travel through the laser detection point) was used as a parameter to determine whether multiple particles are encapsulated in the same droplet. If the peaks
are detected within one unit of the transit time, they are considered in the same droplet. In the interrogation channel, the distance between two droplets was significantly longer than the droplet length. The counter registers the number of the particles in a droplet.

We used our microfluidic sorter to specifically sort two-particle droplets out of a mixed population. Fig. 6.6 shows that such sorting could be successfully applied to droplets containing fluorescent beads or stained HeLa cells. The two-bead/two-cell droplets were sorted into the collection channel, while all other types of droplets (i.e. droplets containing 0, 1, 3, or 4 particles) were directed into the waste channel at a throughput of ~10 droplets/s. The videos that recorded such sorting are included as ESI Videos 2 and 3. We used the images taken by the CCD camera to independently verify the sorting efficiency (Note: the use of the camera is not required for sorting) and the results are shown in Table 5.1. The two-bead droplets and two-cell droplets from a mixed population could be highly enriched after a single pass through the sorter. Whereas the initial percentage of two-bead droplets in the starting population was 5%, the collection channel ended up having 96% two-bead droplets and the waste had only 1% two-bead droplets. Similarly, two-cell droplets were enriched from an initial concentration of 6% to 93% in the collection reservoir. The errors could be attributed to the fact that beads or cells might stick to or overlay on each other while passing the detection point. In principle, further decrease in the size of the interrogation channel could increase the accuracy. However, the practice may also increase the risk of clogging. Thus, a balance between the sorting accuracy and the system robustness needs to be reached.
Fig. 6. Camera images show that only two-particle droplets were sorted out from a mixed population. The droplets containing 0, 1, 3 particles flow into the waste channel (left) and droplets containing (a) two beads or (b) two cells are deflected into the collection channel (right). The droplets were generated with the flow rates 0.15 μl/min for the aqueous phase and 3.60 μl/min for the oil phase. (c) Uniformly two-bead droplets in the outlet reservoir of the collection channel. The droplets were generated under the flow...
rates of 0.27 μl/min for the aqueous phase and 2.10 μl/min for the oil phase. The scale bars represent 100 μm.

<table>
<thead>
<tr>
<th>Particle occupancy per droplet</th>
<th>Bead-containing droplets</th>
<th></th>
<th>Cell-containing droplets</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Input</td>
<td>Collection</td>
<td>Input</td>
<td>Collection</td>
</tr>
<tr>
<td>Number</td>
<td>Percentage (%)</td>
<td>Number</td>
<td>Percentage (%)</td>
<td>Number</td>
</tr>
<tr>
<td>0</td>
<td>1459</td>
<td>68.4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>560</td>
<td>26.2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>101</td>
<td>4.7</td>
<td>79</td>
<td>96.3</td>
</tr>
<tr>
<td>3</td>
<td>12</td>
<td>0.6</td>
<td>3</td>
<td>3.7</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**Table 6.1 Results of sorting out two-bead and two-cell droplets from mixed droplet populations.** The droplets were generated under the flow rate of 0.15 μl/min for the aqueous phase and 3.6 μl/min for the oil phase. The data were extracted from camera images.
Co-culture of heterotypic cells are fundamental for examining cell-cell interactions. Conventional studies of cell-cell interactions are conducted at the population level and do not reveal details at the single cell resolution. Recent advances in micromanipulation technology, such as hydrodynamic trapping, dielectrophoresis, optical tweezers and magnetic force enable scientists to probe cell-cell interaction at the single cell level. However, cell-pairs created by these methods lack isolated environment to avoid cross contamination between different cell pairs. Additionally, most of the current methods can only create and interrogate a low number of cell-pairs, limited by their use for population description.

To address this limitation, we developed a microfluidic droplet based approach in which heterotypical cell pair was isolated inside the same droplet. The previously described droplet sorter has only single signal channel (i.e. single laser and single PMT). In order to differentiate two types of cells with single channel, we develop an approach in which different cell types are stained with the same type of dye but with different fluorescent intensity (Fig.5.7). As a model system for our studies, we use two different concentration (1µM and 50µM) of a green (FITC) fluorescent nucleic acid stain (Syto13, Invitrogen) to label the nucleus of two different population of live cells(type A and type B). By correlating the cell types with fluorescent intensity of their signals, we can differentiate the cell types using single signal channel. For the convenience of observation using microscopic imaging, we deliberately labeled the cytosol of type A cell with Calcein AM red dye, which does not interfere with the FITC nucleus dye. As corroborated by the fluorescence microscopic image, type A cells and type B cells have significantly different intensity of the FITC fluorescence in their nucleus.
**Fig. 6.7 Schematic of differential labeling strategy.** Type A and type B cells are stained with the same type but different concentrated nucleus dye (Sytox 13). The cytosol of type A cells was also stained using Calcein AM Red dye for convenient observation under the microscope. After this treatment, the laser-induced-fluorescence of cells between different populations can be dramatically different.

In order to tell the intensity difference between different signal by our detection system, the previously described sorting system is slightly modified in both the hardware (Fig.5.8) and software parts(Fig.5.9). First of all, the sensitivity for detecting cells in droplets can be increased by using high magnification objectives. Instead of using 10x dry objective(used to differentiate cell number inside droplet), we use 20x dry objective to
detect the intensity of different signal with high sensitivity. Secondly, we use a cylindrical lens to convert the original laser spot into a laser line to promote the uniformity of the laser intensity in the detection area. Additionally, we replace the comparator with an AD converter, which is able to convert the analog signal of detected fluorescent signal into digitalized signal. After AD conversion, the peak value of each cell signal can be obtained, quantified and compared using the microcontroller. Thus, both cell number (number of peaks) as well as cell types (height of peaks) in each droplet can be determined by the microcontroller, which can make sorting decision accordingly in real time.

**Fig. 6.8 Schematic of the modified droplet sorting system for sorting out heterotypic cell pair encapsulated droplet.** An AD converter is integrated into the sorting system to obtain the peak information of each detected cell signal. The sensitivity
of this system is greatly promoted by using objective with higher magnification (20x) and a laser line shaped by a cylindrical lens.

In order to sort droplets based on both the number and types of encapsulated cells, we first examined whether our system generated reliable signal that would indicate both the and type and number of particles in each droplet. As shown in the zoomed-in spectrum in Fig. 6.9, typical fluorescence signal detected from a mixed droplet population with encapsulated fluorescent cells presents peaks of various intensities. Each droplet generates a broad peak due to the light scatter at the water/oil interface. The broad peaks may have one or two embedded high-intensity peaks that were produced by fluorescent cells encapsulated. The variance in the peak intensity was due to the aforementioned differential labeling strategy: high-spiked peaks indicate type B cells, which were stained using high-concentrated nucleus dye; the low-intensity peak indicate type A cells, which were labeled with nucleus dye of low concentration.

Fig. 6.10 shows the algorithm used to sort specifically droplets that contain heterotypic cell pair. The algorithm determines two important facts: First, whether multiple cells are within the same droplet based on the interval between the cells. When the cells are in the same droplet, the interval is shorter than the time required for a droplet to pass through the detection point, assuming that the distance between droplets was significantly longer than the droplet length. The algorithm also included a counter to register the number of particles in one droplet. Second, what type of cells are encapsulated inside the droplet. This can be done by comparing the intensity of cell signals with 3 established threshold, as shown in Fig. 6.9. If detected cell signal is higher than threshold 1, it is classed as type B cell. If the detected signal is higher than threshold 1 but lower than
threshold 2, it is classed as type A cell. Threshold 3 is slightly higher than the signal intensity of empty droplet.

**Fig. 6.9** The fluorescence signal contains information on the type and number of encapsulated cells per droplet. The fluorescence intensity data over time are shown for droplets with no cell, single type B cell and heterotypic cell pair from the left side to the right side in the Fig.. Three different intensity thresholds were established in order to differentiate the cell type and cell number in each individual droplet. The flow rates were set at 0.15 μl/min and 3.6 μl/min for the aqueous phase and the oil phase, respectively.
Fig. 6. 10 A flowchart that summarizes the logic of the microcontroller for sorting heterotypic cell-pair droplets out of a mixed population. AD converter are needed to determine the intensity of each cell signal. The droplet transit time (i.e. the time required for a droplet to travel through the laser detection point) was used as a parameter to determine whether multiple cells are encapsulated in the same droplet. If the peaks are detected within one unit of the transit time, they are considered in the same droplet. In the interrogation channel, the distance between two droplets was significantly longer than the droplet length. The counter registers the number of the cells in a droplet.
Taking advantage of both the novel cell labeling and sorting strategy, we used our microfluidic sorter to specifically sort heterotypic cell-pair droplets out of a mixed population, using only single laser and single PMT detector. The heterotypic cell-pair droplets were sorted into the collection channel, while all other types of droplets (i.e. with other combinations) were directed into the waste channel at a throughput of ~10 droplets/s (Fig. 6.11a). The success of isolation can be further confirmed using fluorescence microscopic image taken at the outlet of the collection channel (Fig. 5.11b).

**Fig. 6.11** Camera images confirm the successful isolation of heterotypic cell-pair from a mixed droplet population. (A) Image show that only two-cell droplets were sorted out from a mixed population. (B) Fluorescence image confirm the heterotypic cell pair in each sorted out droplet. The droplets were generated under the flow rates of 0.27 μl/min for the aqueous phase and 2.10 μl/min for the oil phase. The scale bars represent 100 μm.
6.4 Conclusions

Droplet microfluidics provides a high-throughput platform for screening subjects and conditions involved in biology. Droplets with encapsulated beads and cells have been increasingly used for studying molecular and cellular biology. Droplet sorting is needed to isolate and analyze the subject of interest during such screening. The vast majority of current sorting techniques use fluorescence intensity emitted by each droplet as the only criterion. However, due to the randomness and imperfection in the encapsulation process, typically a mixed population of droplets with an uneven number of encapsulated particles is resulted and used for screening. Thus droplet sorting based on the number and type of encapsulated particles becomes necessary for isolating or enriching droplets with a specific occupancy. In this work, we developed a fluorescence-activated microfluidic droplet sorter that integrated a simple deflection mechanism based on the use of a solenoid valve and a sophisticated signal processing system with a microcontroller as the core. By passing droplets through a narrow interrogation channel, the encapsulated particles were detected individually. The microcontroller conducted the computation to determine the number and type of encapsulated particles in each droplet and made the sorting decision accordingly that led to actuation of the solenoid valve. We tested both fluorescent beads and stained cells and our results showed high efficiency and accuracy for sorting and enrichment.

REFERENCES:


APPENDICES

Appendix A: Procedure for Microfluidic Chip Fabrication (Soft lithography)

*Single Layer:*

Microfluidic devices of single layer are fabricated based on PDMS using standard soft lithography method described below. The microscale patterns are first created using a computer-aided design software (FreeHand MX, Macromedia) and then printed out on high-resolution (5080 dpi) transparencies. The transparencies are used as photomasks in photolithography on a negative photoresist (SU-8 2025, MicroChem Corp.). The thickness of the photoresist and hence the depth of the channels is determined by the spin rate during spin-coating. The pattern of channels in the photomask is replicated in SU-8 after exposure and development. The microfluidic channels are molded by casting a layer (~5 mm) of PDMS prepolymer mixture (General Electric Silicones RTV 615, MG chemicals) with a mass ratio of A:B = 10:1 on the photoresist/silicon wafer master treated with tridecafluoro-1,1,2,2-tetrahydrooctyl-1-trichlorosilane (United Chemical Technologies). The prepolymer mixture is cured at 85°C for 2 hours in an oven and then peeled off from the master. Glass slides are cleaned in a basic solution (H2O: NH4OH (27%) : H2O2 (30%) = 5:1:1, volumetric ratio) at 75°C for an hour and then rinsed with DI water and blown dry. The PDMS chip and a glass slide are rendered hydrophilic by oxidizing them using a plasma cleaner in atmosphere. The PDMS chip is then immediately brought into contact against the slide after
oxidation to form closed channels.

**Double Layer with Microfluidic Valve.**

The fabrication of device with microvalve is followed by multilayer soft lithography with substantial modifications. The control layer master is made using a negative photoresist SU-8 2075 (Microchem) with a thickness of 50 µm. The fluidic layer master is made using a negative photoresist (Microchem) with a thickness of 40 µm. The thickness of the photoresist is translated into the depth of the microfluidic channel. The control layer of the device has a PDMS composition (GE Silicones RTV 615, MG Chemicals, mass ratio of A:B=20:1). The fluidic layer of the device has a PDMS composition (GE Silicones RTV 615, MG Chemicals, mass ratio of A:B=5:1). The control layer and the fluidic layer are pre-heated for about 30 min to solidify both layers. The thickness of fluidic layer depends on spin rate of liquid PDMS. The control layer had a thickness ~0.5 cm. The two layers are bonded together upon contact followed by hole-puncturing. Then the combined layers are heated for another 2 hours. Then the combined layers are bonded to a clean glass slide using the plasma cleaning method as described in the single layer method. Finally the entire device is baked in an oven at 80°C for 2 hr.
Appendix B: Cell Culture

**Chinese Hamster Ovary Cells.**

Chinese hamster ovary (CHO-K1) are cultured in plastic tissue culture flasks at 37°C, under 5% CO2 in the Dulbecco’s modified Eagle’s medium (DMEM, Mediatech Inc.) supplemented with 10% (v/v) fetal bovine serum (FBS, Sigma), penicillin (100 units/ml, Sigma), and streptomycin (100µg/ml, Sigma). The cells are trypsinized and diluted at a ratio of 1:5 every 2 days to maintain them in the exponential growth phase (~1×10^6 cells/ml).

**Hela(CCL-2) cells.**

HeLa (CCL-2) cells were grown at 37°C with 5% CO2 in Dulbecco’s modified Eagle’s medium (DMEM) (Mediatech, Herndon, VA) supplemented with 10% (v/v) fetal bovine serum (Sigma) and 1% penicillin (100 mg/ml, Sigma). The cells were trypsinized and diluted at a ratio of 1: 5–1: 8 every 2 d to maintain the cells in the exponential growth phase. The harvested cells were centrifuged at 300 g for 5 min and resuspended in DMEM culture medium at a final concentration of 1×10^6 cells/ml before experiments.

**GM12878 cells.**

GM12878 cells were obtained from Coriel Institute for Medical Research.
Species of origin of the cell line was confirmed by PCR targeting the glucose-6-phosphate dehydrogenase gene. Donor subject has a single bp (G-to-A) transition at nucleotide 681 in exon 5 of the CYP2C19 gene (CYP2C19*2) which creates an aberrant splice site. Donor origin of the cell line was confirmed using PCR against the point mutation. The cell line was tested for mycoplasma contamination using ABI MycoSEQ mycoplasma detection assay (Applied Biosystems). Cells were propagated in RPMI 1640 (11875-093, Gibco) plus 15% fetal bovine serum (26140-079, Gibco), 100 U penicillin (15140-122, Gibco), and 100µg/ml streptomycin (15140-122, Gibco) at 37°C in a humidified incubator containing 5% CO2. Cells were sub-cultured every two days to maintain them in exponential growth phase.

**Chinese Hamster Ovary Cells with p65-GFP Expression.**

When the cells reach ~90% confluency in a 75 cm flask, they are ready for subculture. The culture media are removed from cells by aspiration, followed by rinsing with PBS and aspiration again. 2 ml of 0.25% trypsin/0.53 mM Tris-EDTA solution is added to the culture dish and incubated for 2~3 min at room temperature. Then 10 ml of pre-warmed complete growth media (Hams F12K supplemented with 10% fetal bovine serum, 100 units/ml penicillin, 100 µg/ml streptomycin, 100 µg/ml hygromycin B) is added and pipette up and down to break up any clumps. Detached cells are transferred to a 15 ml centrifuge tube and centrifuged at 125 g for 5 min to collect the cells, which is resuspended in a 10 ml complete growth media. 2 ml of suspended cells are dispensed into the same flask containing 13 ml complete growth media, which is placed in a humidified incubator at 37 °C
with 5% CO2.

Appendix C Procedure for Quantitative Detection of Nucleocytoplasmic Transport

Materials

Cell culture and sample preparation

1. Hela cells (ATCC, CCL-2)
2. Cell culture medium: Dulbecco’s modified Eagle’s medium (DMEM) supplemented with 10% (vol/vol) FBS, 2 mM l-glutamine, 100 U ml $^{-1}$ penicillin and 100 mg ml $^{-1}$ streptomycin. Store at 4 °C for up to 1 year.
3. Fetal bovine serum (FBS)
4. Penicillin-streptomycin solution (Invitrogen)
5. l-glutamine solution (Invitrogen)
6. Trypsin-EDTA (0.05% (wt/vol); Invitrogen)
7. Sterile Phosphate-buffered saline (PBS): Dissolve one PBS buffer tablet (Sigma-Aldrich) in 100ml DI water. Autoclave for 20 min to sterilize. Store at 4 °C for up to 1 year.
8. Recombinant Human TNF alpha (AbD Serotec): Aliquot into smaller volume (e.g. ~10μl) to avoid extra freeze-thaw cycles.

Immunofluorescence labeling
1. Blocking solution: Prepare a solution of 0.5% Bovine Serum Albumin (BSA) (wt/vol) in sterile PBS buffer. Filter-sterilize with a 0.2-µm membrane and store at 4 °C for up to 2 months.

2. Fixation solution: Dilute 16% formaldehyde (FA) in sterile blocking buffer to a final concentration of 4%. Filter-sterilize with a 0.2-µm membrane and store at 4 °C for up to 2 months.

3. Permeabilization solution: Dilute 1% Triton X-100 in blocking buffer to a final concentration of 0.2%. Filter-sterilize with a 0.2-µm membrane and store at 4 °C for up to 2 months.

4. Selective releasing solution: Dissolve 0.05% (w/v) Saponin (Chem Service) in Dulbecco’s modified Eagle’s medium (DMEM). Filter-sterilize with a 0.2-µm membrane and store at 4 °C for up to 1 months.

5. Primary antibody: NF-κB p65 (Santa Cruz Biotechnology)

6. Secondary antibody: DyLight™ 488 Goat anti-mouse IgG1 antibody (Biolegend)

7. Sodium azide

**Equipment**

**Solution preparation**

1. Milli-Q Direct system (Millipore)
2. Glass syringe with Lure-Lock tip (50 ml)
3. Sterile syringe filters with 0.2 µm pore diameter
**Cell culture and sample preparation**

1. Laminar flow hood
2. Cell culture incubator set at 37 °C with humidified, 5% CO₂ atmosphere
3. Cell culture flasks and well plates
4. Pipettes/tips
5. Centrifuge tubes (0.6, 1.5, 15 and 50 ml)
6. Centrifuge
7. Spray bottles
8. Hemocytometer
9. Ice maker
10. Mini lab rotator (24 rpm)
11. 30 μm Nylon-mesh Filters (Partec)

**Microscope**

1. Inverted phase-contrast and epifluorescence microscope system with 100 W mercury lamp (IX-71; Olympus) equipped with ×10 and ×20 objectives (0.5 NA)
2. High-resolution CCD camera (Hamamatsu, ORCA-285)
3. Fluorescence filter cubes (exciter HQ480/40x, emitter HQ535/50m; Chroma Technology)
4. Image acquisition software (QImaging QCapture Pro)
5. ImageJ software (or similar)
Flow Cytometer

1. BD FACSCanto II workstation (BD Bioscience)
2. Sapphire 488-20 laser (488nm, 20mW, Coherent)
3. BP Filter or LP mirror (530/30, intended dye: FITC)
4. Flow cytometer controlling and analysis software (BD FACSDiva)

Methods

Cell sample preparation

1. Use cells of 80% confluency for the analysis. For Hela cells, this corresponds to about 3.0 × 10^6 cells per 75 cm² flask. This total amount of cells is sufficient for 5~6 samples (given that ideally 5× 10^6 cells are needed for one flow cytometric screening) (see Note 1).
2. Aspirate the medium and briefly rinse the adherent cells with 10 ml of ice-cold PBS.
3. Aspirate PBS and add 3 ml of the trypsin-EDTA. Disperse the solution by gently shaking the plate and incubate it in the incubator for 5 min at 37°C (see Note 2).
4. Gently shake the plate until cells detach. Detached cells can be easily identified under a microscope, as they come off in round shape.
5. Immediately add 10 ml of cell culture medium to inactivate trypsin. Disperse suspended cells in medium by gently pipetting three or four times to produce a single-cell suspension (see Note 3).
6. Transfer the suspension to a 15-ml conical tube and pellet the cells by centrifugation at 300g for 5 min at 4 °C (see Note 4).

7. Discard the supernatant and resuspend the cell pellet in 1.5 ml of fresh cell culture medium by gently pipetting with 1ml pipet tip several times. Put the cells on ice and use a small aliquot to estimate the number of cells using a hemocytometer (see Note 5).

8. Aliquot around one-sixth of the cell suspension (250 µl and ~5× 10⁵ cells) to a new 0.6 ml centrifuge tube (see Note 6).

**Cell stimulation**

1. Add TNF-α into the cell suspension and mix the solution by gently pipetting up and down several times.

2. Stimulate cells by incubation with TNF-α for a period of time at 37 °C under constant, mild agitation (>30 min for complete NF-κB translocation from the cytosol to nucleus).

3. Centrifuge the cell suspension at 300g for 5 min at 4 °C.

4. Slowly aspirate and discard the supernatant. Add 200µl of the fresh DMEM, and gently disperse the pellet by pipetting up and down until the cells are resuspended.

5. Repeat steps 11 and 12 once.

**Selective release of cytosolic proteins**

1. Centrifuge the cell suspension at 300g for 5 min at 4 °C.

2. Slowly aspirate and discard the supernatant and resuspend cells with 200 µl of the selective releasing solution by gently pipetting several times.
3. Incubate cells with the selective releasing solution for 5 min on a rotator mixer (24 rpm) at the room temperature (see Note 7 and 8).

4. Centrifuge the cell suspension at 600g for 5 min at 4 °C (see Note 9).

5. Slowly aspirate and discard the supernatant and resuspend cells with 200 μl of the blocking solution by gently pipetting several times.

6. Repeat steps 17 and 18 once.

**Fixation and permeabilization of treated cells**

1. Centrifuge the cell suspension at 600g for 2 min at 4 °C.

2. Slowly aspirate and discard the supernatant and resuspend cells in 200 μl pre-warmed fixation solution by gently pipetting several times (see Note 10).

3. Incubate cells with the fixation solution for 40 min at room temperature (see Note 11).

4. Centrifuge the cell suspension at 600g for 5 min at 4 °C.

5. Slowly aspirate and discard the supernatant and wash the sample with 200 μl blocking solution by gently pipetting several times.

6. Repeat steps 23 and 24 once.

7. Centrifuge the cell suspension at 600g for 5 min at 4 °C.

8. Slowly aspirate and discard the supernatant and resuspend the cells in 200 μl of the permeabilization solution by gently pipetting several times.

9. Incubate cells in the permeabilization solution for 5 min at room temperature (see Note 12).

10. Centrifuge the cell suspension at 600g for 5 min at 4 °C.
11. Slowly aspirate and discard the supernatant and wash the cells with 200 μl of the blocking solution by gently pipetting several times.

12. Repeat steps 29 and 30 once.

**Fluorescence immunostaining**

1. Add 0.75 μl of the primary antibody solution (1:200 diluted, containing ~100 ng antibody) to the solution and mix by pipetting several times (see Note 13).

2. Incubate cells with the primary antibody for 1 h at room temperature.

3. Centrifuge the cell suspension at 600g for 5 min at 4 °C.

4. Slowly aspirate and discard the supernatant and wash the cell pellet with 200 μl of the blocking solution by gently pipetting several times.

5. Repeat steps 34 and 35 once.

6. Centrifuge the cell suspension at 600g for 5 min at 4 °C.

7. Slowly aspirate and discard the supernatant and resuspend the cells in 200 μl blocking solution by gently pipetting several times.

8. Add 0.5 μl (containing ~50 ng) of fluorophore-conjugated secondary antibody (1:300 diluted) and mix the solution by pipetting several times (see Note 14).

9. Incubate cells with the secondary antibody for 30 min at the room temperature in the dark.

10. Centrifuge the cell suspension at 600g for 5 min at 4 °C.

11. Slowly aspirate and discard the supernatant and wash the cell pellets with 200 μl blocking solution by gently pipetting several times.

12. Repeat steps 41 and 42 once.
13. Centrifuge the cell suspension at 600g for 5 min at 4 °C.

14. Slowly aspirate and discard the supernatant and resuspend the pellet with 200 μl PBS solution containing 0.1% sodium azide by gently pipetting several times.

15. The cells can be stored at 4°C for up to 24 h, if not immediately analyzed by flow cytometry.

**Observation by fluorescence microscopy (optional)**

1. At this stage of the protocol, it is optional to use an aliquot of the prepared cell sample for analysis by fluorescence microscopy. The images confirm the proposed mechanism. Our technique renders the fluorescence intensity different for cells with translocation and those without translocation, whereas standard immunostaining reveals the different localizations of the protein via imaging with the overall fluorescence intensity from whole cells being the same for the two types of cells (see [Note 15](#)).

   i. Transfer immunostained cell samples to 96 well plate. Make sure each well has around 1,000 cells and 200 μl PBS solution.

   ii. Mix the cell solution by pipetting several times in each well.

   iii. Centrifuge the sample at 50g for 15 min at 4°C so that the floating cells settle at the bottom of wells.

   iv. Capture both phase-contrast and fluorescent images of cells at 5 different locations in each sample using a fluorescence microscope equipped with a ×20 dry objective and a CCD camera. The excitation and emission were filtered by a fluorescence filter cube for observing green fluorescence.
Cell screening by a common flow cytometer

1. Start the flow cytometer and the computer, perform the usual fluidics maintenance and launch FACSDiva 6.0.

2. Open the appropriate acquisition document from the storage folder or create a new one.

3. Remove aggregated cells from the sample using a nylon-mesh filter (see Note 16).

4. Run samples on the flow cytometer with 488 nm excitation.

5. When acquiring data on the flow cytometer, create a forward scatter (FSC) versus side scatter (SSC) plot. To exclude cell debris, air bubbles and laser noise from analysis, ensure that the expected cell population is visible in FSC/SSC plot by adjusting the individual FSC (voltage 200-400, linear mode) and SSC (voltage 200-400, linear mode) settings.

6. A gate is applied to the region around your cell population of interest on this FSC /SSC plot.

7. Necessary fluorophore analysis histogram/plots can now be created as needed (i.e. FSC vs counts or FSC vs FITC).

8. Collect at least 20,000 events in each sample.

9. After completing data acquisition, perform the usual fluidics maintenance and proceed to data analysis with Flow Jo (version 7.6.1).
Notes

1. Each cell line has a characteristic optimal confluency and should be regularly sub-cultured. Hela cells are passaged every 2~3 days at a ratio of 1:10 by trypsinization.

2. The trypsinization time may need to be optimized for a particular cell type. Insufficient trypsinization may lead to cell aggregates. Excessive trypsinization may cause cell damage or death.


4. If multiple culture flasks are used, mix all cell suspensions in one tube for cell counting to avoid variability among different cultures.

5. High shear stress induced by vigorous pipetting may damage cells and cause nucleocytoplasmic translocation to create artifacts in the final result.

6. Single cell suspension must be obtained at this point for optimal TNF-α stimulation and selective releasing.

7. The concentration of saponin in the selective releasing solution should be optimized for different cell/target protein combinations. Low saponin concentration may lead to incomplete releasing of cytosolic proteins. High concentration of saponin may cause releasing of nuclear proteins. The cytosolic proteins NF-κB of Hela cells can be completely released by selective releasing solution with 0.05% saponin in 5 min. The saponin concentration and the treatment time need to be optimized together.

8. The treatment time for the selective releasing should be optimized for different cell/target protein combinations. Insufficient incubation time may lead to incomplete releasing of cytosolic proteins. Prolonged incubation time may cause
rupture of nuclei and releasing of nuclear proteins. In our experiment, the cytosolic NF-κB of Hela cells can be completely released in 5 min by the selective releasing solution (containing 0.05% saponin).

9. The density of cells decreases after the selective releasing. Increased centrifugation speed is critical to pellet most of the cells.

10. Pre-warming the fixation solution at 37 °C is critical for maintaining cell integrity.

11. Cells after selective release are easy to rupture. A prolonged fixation time of >30 min instead of the standard 10 min treatment is critical for the integrity of these cells.

12. The permeabilization provides access to intracellular target proteins. Instead of the standard 15 min permeabilization time recommended in most conventional immunostaining protocols, Hela cells after the selective release can be completely permeabilized by Triton X-100 within 5 min. The permeabilization time ought to be optimized for a particular cell type. Insufficient permeabilization leads to inadequate labeling. Prolonged permeabilization may cause loss of the target protein and decreased cell integrity.

13. The amount of the primary antibody needs to be optimized. Insufficient antibody leads to incomplete labeling. Excessive antibody leads to increased non-specific binding. In our experiment, 200 ng primary antibody per $1 \times 10^6$ cells yields the best result. It is also crucial to keep staining conditions (i.e. concentrations of cells and antibodies) consistent among samples in order to generate good reproducibility.
14. We used Alexa Fluor 488 conjugates, with excitation and emission properties nearly identical to those of FITC. Compared to FITC, Alexa Fluor 488 is brighter for flow cytometry and fluorescence microscopy.

15. Use the centrifugation speed of 50g to avoid concentration of cells at the edge of the well bottom.

16. Cell aggregates potentially clog the tubing of the flow cytometer

<table>
<thead>
<tr>
<th>Step</th>
<th>Problem</th>
<th>Possible reason</th>
<th>Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 3 of 3.1</td>
<td>Cell clumping</td>
<td>Insufficient trypsinization</td>
<td>Increase the amount of trypsin-EDTA solution;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Increase the time of trypsin-EDTA incubation</td>
</tr>
<tr>
<td>Step 3 of 3.1</td>
<td>Reduced cell viability</td>
<td>Excessive trypsinization</td>
<td>decrease the amount of trypsin-EDTA solution;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>decrease the time of trypsin-EDTA incubation</td>
</tr>
<tr>
<td>Step 1 of 3.2</td>
<td>No translocation</td>
<td>Reagent out of date; Insufficient mixing</td>
<td>The shelf life of human recombinant TNFα is 2 month at -20 °C. Agitate the cell solution several times during stimulation.</td>
</tr>
<tr>
<td>Step 4 of 3.3</td>
<td>Excessive loss of cells after selective releasing</td>
<td>Inefficient pelleting</td>
<td>Increase the centrifugal force above 600g</td>
</tr>
<tr>
<td>Step 4 of 3.4</td>
<td>Loss of fixed cells</td>
<td>Inefficient pelleting</td>
<td>Increase the centrifugal force above 600g</td>
</tr>
<tr>
<td>Step 14 of 3.5</td>
<td>Loss of cells during immunostaining</td>
<td>cell absorption on the tube surface;</td>
<td>Use clear, low-retention tubes;</td>
</tr>
</tbody>
</table>
cell rupture due to excessive mechanical force

| Step 5 of 3.7 | Excessive amount of outliers | Too few cells are examined | Eliminate subsets in the FSC vs. SSC plot which contain <100 cells; Acquire a minimum amount of 20,000 exponentially growing cells for analysis |

*Table 1 of Appendix C Troubleshooting*
## Appendix D: Spectrum of Fluorophores

<table>
<thead>
<tr>
<th>Fluorophore</th>
<th>EX</th>
<th>EM</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hoechst</td>
<td>350</td>
<td>461</td>
<td>cell-permeant nuclear counterstain</td>
</tr>
<tr>
<td>GFP</td>
<td>478</td>
<td>507</td>
<td>protein labeling in live cells</td>
</tr>
<tr>
<td>DIOC18(3)</td>
<td>484</td>
<td>501</td>
<td>membrane stain</td>
</tr>
<tr>
<td>EGFP</td>
<td>487</td>
<td>509</td>
<td>protein labeling in live cells</td>
</tr>
<tr>
<td>Syto 16</td>
<td>488</td>
<td>518</td>
<td>nucleic acid stain; permeant to live</td>
</tr>
<tr>
<td>YOYO-1</td>
<td>491</td>
<td>509</td>
<td>nucleic acid stain; impermeant to</td>
</tr>
<tr>
<td>Calcein</td>
<td>494</td>
<td>514</td>
<td>to determine cell viability</td>
</tr>
<tr>
<td>Sytox Green</td>
<td>504</td>
<td>523</td>
<td>nucleic acid stain; impermeant to live cells</td>
</tr>
</tbody>
</table>

*EX*: excitation wavelength; data from Invitrogen, Inc.

**EM**: emission wavelength; data from Invitrogen, Inc.
Appendix E: MCU code for droplet sorting based on number of cell encapsulated in one droplet

interval EQU 00H
jet EQU 01H
cell_number EQU 02H

ORG 0000H
AJMP MAIN

ORG 0003H
AJMP INT_0

ORG 000BH
AJMP TIME0

ORG 30H
MAIN:
MOV SP,#5FH
CLR IT0
MOV TMOD,#01H
MOV TH0,#0F5H
MOV TL0,#33H
SETB EA
SETB ET0
SETB EX0
SETB IT0
CLR P2.5
MOV interval,#1
MOV jet, #3
MOV cell_number, #0

;---------------------------------------
CELL_CHECK:

AJMP CELL_CHECK

;---------------------------------------
TIME0:
PUSH ACC
PUSH PSW
CLR TR0
CLR EX0
TRET: MOV TH0,#0F5H
MOV TL0,#33H

CELL_JUDGE:
MOV a, cell_number
MOV cell_number,#0
CJNE a, #02, FINISH

;-----------------------------------
AFTER_DELAY:
MOV R7,interval
A1: MOV R6,#10
A2:
DJNZ R6,A2
DJNZ R7,A1

VALVE:
SETB P2.5

;---------------------------------------
JET_DELAY:
MOV R5,jet
J1: ACALL D1MS
DJNZ R5,J1

;----------------------------------------
CLR P2.5
FINISH:
SETB EX0
POP PSW
POP ACC
RETI

INT_0:
PUSH ACC
PUSH PSW
SETB TR0
INC cell_number
POP PSW
POP ACC
RETI

D10MS: MOV R7,#50
   D1:MOV R6,#100
   D2:DJNZ R6,D2
   DJNZ R7,D1
   RET

D1MS: MOV R7,#3
   D1:MOV R6,#100
   D2:DJNZ R6,D2
   DJNZ R7,D1
   RET

END
Appendix F: MCU code for droplet sorting based on number and type of cell encapsulated in one droplet

Without using timer:

```c
#include<reg51.h>
#include<intrins.h>

sbit ADC_CS = P3^4
sbit ADC_CLK = P1^0
sbit ADC_DI = P1^1
sbit ACC0 = ACC^0
sbit valve = P1^5
sbit id1 = P1^4
sbit id2 = P1^3

unsigned char typeA = 0x00;
unsigned char typeB = 0x00;
unsigned char count = 0x00;

void delayms(unsigned char ms)
```
{ 
    unsigned char i;
    while(ms--)
    {
        for(i = 0; i < 122; i++);  
    }
}

void indicator1(void)
{
    id1=1;
    delayms(2);
    id1=0;
}

void indicator2(void)
{
    id2=1;
    delayms(2);
    id2=0;
}

ADC_read()
{
    unsigned char i;
    ADC_CS=1;
    ADC_CLK=0;
    ADC_CS=0;
ADC_DI=1;
_nop_();

ADC_CLK=1;
ADC_DI=0;
ADC_CLK=0;
ADC_DI=1;
_nop_();

ADC_CLK=1;
ADC_DI=0;
ADC_CLK=0;
ADC_DI=0;
_nop_();

ADC_CLK=1;
ADC_DI=1;
ADC_CLK=0;
_nop_();

ADC_CLK=1;
ADC_DI=1;
ADC_CLK=0;

ACC=0;
for(i=8;i>0;i--)
{
ADC_CLK=0;
ACC=ACC<<1;
ACC0=ADC_DI;
ADC_CLK=1;
}
ADC_CS=1;
return(ACC);

void INT_init (void){
    EA = 1;
    EX0 = 1;
    IT0 = 1;
}

void Timer0_init (void){
    TMOD = 0x01;
    ET0 = 1;
    TH0 = 0x3C;
    TL0 =0xB0 ;
    count=20;
}
void Valve_init(void){
    valve=0;
}

unsigned char flag = 0;
void ExtInt0(void) interrupt 0
{
    EX0 = 0;
    flag = 1;
    TR0=1;
}

void timer0(void) interrupt 1
{
    EX0 = 0;
    Timer_id=1;
    if (count==1)
    {
        Timer_id=0;
        TR0 = 0;
        count=20;
    }
if (typeA == 2 && typeB == 0)
{
    valve = 1;
    delayms(10);
    valve = 0;
}
typeA = 0x00;
typeB = 0x00;
}

else
{
    count --;
}
TH0 = 0x3C;
TL0 = 0xB0;
EX0 = 1;
}

main()
{

Valve_init();
INT_init();
Timer0_init();

while(1){

    if(flag == 1){
        unsigned char array[3] ={0};
        unsigned char i1,i2,max;

        array[0]=ADC_read();
        _nop_();
        _nop_();
        _nop_();
        _nop_();
        array[1]=ADC_read();
        _nop_();
        _nop_();
        _nop_();
        _nop_();
        array[2]=ADC_read();
        _nop_();
        _nop_();
        _nop_();
        _nop_();
        while(1){
            i1=(array[0]/3+array[1]/3+array[2]/3);
i2=ADC_read();
_nop_();
_nop_();
_nop_();
_nop_();
i2=(i2/3+array[1]/3+array[2]/3);

if(i1>i2){
    max = i1;
    break;
}
Else{
    array[0]=array[1];
    array[1]=array[2];
    array[2]=i2;
}
}

if (max>=0xBC)
{
    typeA++;
}
else
{
    typeB++;
}
indicator1();

}

flag=0 ;
EX0 = 1 ;

}
//Using timer

#include<reg51.h>
#include<intrins.h>

sbit ADC_CS = P3^4;
sbit ADC_CLK = P1^0;
sbit ADC_DI = P1^1;
sbit ACC0 = ACC^0;
sbit valve = P1^5;
sbit id1 = P1^4;
sbit id2 = P1^3;

unsigned char typeA = 0x00;
unsigned char typeB = 0x00;
unsigned char count = 0x00;

void delayms(unsigned char ms)
{
    unsigned char i;
    while (ms--)
    {
        for(i = 0; i < 122; i++);
    }
}
void indicator1(void)
{
    id1=1;
    delayms(2);
    id1=0;
}

void indicator2(void)
{
    id2=1;
    delayms(2);
    id2=0;
}

ADC_read()
{
    unsigned char i;
    ADC_CS=1;
    ADC_CLK=0;
    ADC_CS=0;
    ADC_DI=1;
    _nop_();
ADC_CLK=1;
ADC_DI=0;
ADC_CLK=0;
ADC_DI=1;
_nop_();

ADC_CLK=1;
ADC_DI=0;
ADC_CLK=0;
ADC_DI=0;
_nop_();

ADC_CLK=1;
ADC_DI=1;
ADC_CLK=0;
_nop_();

ADC_CLK=1;
ADC_DI=1;
ADC_CLK=0;
_nop_();

ADC_CLK=1;
ACC=0;
for(i=8;i>0;i--)
{
  ACC=ACC<<1;
  ACC0=ADC_DI;
}
ADC_CLK=1;
}
ADC_CS=1;
return(ACC);
}
void INT_init (void){
    EA = 1;
    EX0 = 1;
    IT0 = 1;
}
/*******************************************************************************/
void Timer0_init (void){
    TMOD = 0x01;
    ET0 = 1;
    TH0 = 0x3C;
    TL0 =0xB0 ;
    count=20;
}

void Valve_init(void)
{
    valve=0;
}

unsigned char flag = 0;

void ExtInt0(void) interrupt 0
{
    EX0 = 0;
    flag = 1;
    TR0=1;
}

void timer0(void) interrupt 1
{
    EX0 = 0;

    // Timer_id=1;

if (count==1)
{

    // Timer_id=0;

    TR0 = 0;
    count=20;

    if (typeA ==1 && typeB==1)
    {
        valve=1;
        delayms(10);
        valve=0;
    }
    typeA=0x00;
    typeB=0x00;
}

else
{
    count--;

}

TH0 = 0x3C;

TL0 = 0xB0;
EX0 = 1;

main()
{
  Valve_init();
  INT_init();
  Timer0_init();

  while(1){
    if(flag == 1){
      unsigned char array[3] = {0};
      unsigned char i1, i2, max;

      array[0] = ADC_read();
      _nop_();
      _nop_();
      _nop_();
      _nop_();
      array[1] = ADC_read();
      _nop_();
      _nop_();
      _nop_();
      _nop_();
      array[2] = ADC_read();
      _nop_();
      _nop_();
_nop_();
while(1){
    i1=(array[0]/3+array[1]/3+array[2]/3);
    i2=ADC_read();
    _nop_();
    _nop_();
    _nop_();
    _nop_();
    i2=(i2/3+array[1]/3+array[2]/3);
    if(i1>i2){
        max = i1;
        break;
    } else{
        array[0]=array[1];
        array[1]=array[2];
        array[2]=i2;
    }
}
if (max>=0xBC)
{

typeA++; 

} 

else 

{ 


typeB++; 

indicator1(); 

} 

flag=0 ; 
EX0 = 1 ;