Towards Automated Epigenetics: Sample Processing with Droplet Microfluidics

by

Yi Xu

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemistry)
in the University of Michigan
2018

Doctoral Committee:

Professor Ryan C. Bailey, Chair
Professor Jianping Fu
Professor Robert T. Kennedy
Dr. Tamas Ordog, Mayo Clinic

Yi Xu

yixu@umich.edu

ORCID iD: 0000-0002-4927-8746

Dedication

This dissertation is dedicated to my family, friends, and those who have inspired me to be a better person, especially Loki, my dearest cat.

Acknowledgements

I would like to thank my committee members for their guidance, the National Institutes of Health (NCI CA191186) and the Midwest Cancer Nanotechnology Training Center for the financial support, and the Michigan Center for Integrative Research in Critical Care (MCIRCC) for the patient samples. The work in this dissertation would not be possible without their help and contribution.

I want to express the sincerest appreciation to my advisor, Professor Ryan Bailey, for trusting me to work on a challenging research project that was unfamiliar to me when I joined the group. He has been fully supportive and encouraging, helping me become a better scientist throughout the years. His mentorship will always be with me.

My gratefulness is also with Dr. Tamas Ordog and Dr. Jeong-Heon Lee, my collaborators at Mayo Clinic, Rochester. I have been embracing the beauty of epigenetics because of your time and efforts. Thank you, Dr. Ordog, for serving on my committee and sharing your insights on the promising future of the project; thank you, Dr. Lee, for your kind introduction of me to the members at the Epigenomics Development Laboratory and your mentorship on epigenetics.

My thankfulness to the Bailey lab members is destined to be special and more than because we experienced the lengthy lab translocation and reestablishment as a team. Former lab members Richard Graybill and James Wade, thank you both for listening to my struggles with research and offering valuable suggestions even though epigenetics is out of your expertise. I have been deeply impacted by your independence in research and development. Former lab member and current postdoc Heather Robison, thank you for being a wonderful roommate and a great colleague to

discuss science and personal life with. You as a strong and independent woman always inspire me to live a better life. Former lab member Alex Stanton, thank you for being glamorous with me, going out for brunches on Sundays during the first academic year after we landed in Ann Arbor and shook in the winter. Steven Doonan and Gloria Diaz, thank you for being on the team microfluidics with me to discover the magic of droplets. Gloria, I am very proud to bond with you through so many chic items including anime. Additionally, I enjoy mentoring and working with you for eventually another person involved heavily in biology joined our subgroup to share my love and hate for cells. Former lab member Jamy Lee, you are such a happy person and my meal partner that your excitement at food always soothes me even on the most difficult days. All the postdocs and the younger graduate students in the lab, thank you all being good lab mates of me.

Loki the prince cat, thank you. You have been a great comfort and company for me. Finding you waiting for me at the door is most rewarding when I come home, which erases all the possible negative emotions. Let's be family to each other for all the years upcoming, and you are my dearest baby forever.

Lastly, I thank my family and friends for always being supportive during my chase in education and career. Grandma and grandpa, thank you for raising me up and teaching me the basic principles of being a good person. Mom and dad, thank you for being open-minded and supporting all the decisions I have made for my life. Ran He, Shuqi Lai, Yuwei Tian, and Mengting Wan, you are my wonderful friends sharing the sweet and the bitter of graduate school and always motivate me to advance forward. I am confident that bright future is awaiting, and we will embrace all the good that is to happen on every single tomorrow.

Table of Contents

List of Figures

List of Tables

TABLE

List of Equations

EQUATION

List of Appendices

Abstract

Epigenome regulates gene expression without changing the order of nucleotide sequences in DNA and has been confirmed to be involved in enormous numbers of biological processes including cell proliferation and differentiation, DNA replication and repair, and disease development. It is of medical relevance to further our understanding of the complex network of epigenomic regulation and apply the knowledge in disease diagnosis, prognosis, and therapy development. Eventually, profiling and screening personal epigenome can make a crucial part in precision medicine. Current investigating methods of epigenome usually start with a large number of cells and rely on manual processing of massive numbers of sample preparation prior to data readout, unrealistic in clinical settings where sample sizes may be limited, user qualification may vary, and human resources may also be inefficient. Signals from single cells are also buried due to bulk processing, limiting the application of these powerful assays in heterogeneous epigenome mapping. Thus, automating epigenomic assays using droplet microfluidics were proposed, aiming to utilizing the favorable features of droplets to approach flexible sample-sized, automated epigenomic assays starting directly from cells for qualified clinical usages.

The first trial was using droplet microfluidics to automate samples preparation for an assay probing nucleosome positioning called MNase-seq. Using a droplet microfluidic device to encapsulate cells with lysis and digestion reagent, cellular chromatin was digested extensively to mononucleosomes (yield at ~ 80%) with sample sizes ranging from 125,000 cells to 2,500 cells. Next-generation sequencing of purified mononucleosomal DNA from these samples validated the efficacy of this platform to generate qualified samples for nucleosome mapping with results in

consistency with literatures. Clinical samples from 5 pairs of septic shock patients and matched non-septic controls along with 3 healthy controls (buffy coat extracted directly from whole blood) were also successfully processed by the same strategy and the sequencing results conveyed diagnostic values from the identified nucleosome positioning. More patients need to be enrolled to continue the clinical study to further validate the clinical relevance of nucleosome mapping from samples generated droplet microfluidically.

To push forward the droplet microfluidic sample preparation, another device capable of preparing two samples for two different assays from the same cell sample was designed and characterized. Using a bifurcating channel design, droplets were aliquoted randomly to two directions allowing digestion of chromatin to different extents. Nucleosome-depleted regions as well as nucleosomal regions were both recovered for the detection of accessible regions and nucleosome locations, respectively. qPCR quality control results confirmed the enrichment of nucleosome-depleted, linker regions where regulatory sequences are located from the samples prepared on the device (fold change ≥ 5). Both cell line cells and clinical samples were processed with the optimal design for sequencing to achieve a final validation.

Besides chromatin processing and related assay automation, chromatin immunoprecipitation (ChIP), the gold standard probing DNA-protein interaction, was also transformed to droplet microfluidics. The preliminary results confirmed that the conditions on the droplet microfluidic device were compatible with ChIP. Further optimization was required to improve the enrichment of targeted regions from ChIP DNA collected out of the droplet microfluidic device. The automation level of the entire workflow also needed enhancement by integrating the immunoprecipitating device with the immuno-bead conjugate wash device.

Chapter 1 Introduction

Epigenetics

The biologist Dr. Conrad H. Waddington conceptualized "epigenotype" in the early 1940s and described it as the whole complex of developmental processes connecting the genotype and the phenotype, informally defined the research field of epigenetics as the study of changes in gene function that are heritable and that do not involve any changes in the sequence of DNA.[1-3] Ever since that, epigenetics has developed into one of the most popular research fields. Increasing amount of efforts have been made yielding extraordinary discoveries in elucidating the significant roles epigenetic marks play in the regulation of gene activity, cellular phenotypes, and disease development including cancer.[4] The reversibility and the medical relevance of epigenome have attracted increasing amount of interest over the years, especially the potential of translating epigenetics to precision medicine, one of the most exciting directs in modern healthcare.

Common Epigenetic Marks. The reason for the fact that a single genome gives rise to various cell types within an organism is the mediation of gene expression from the epigenetic marks and regulators to adjust the set of proteins that are expressed and determines the fate of the cell.[4] Many common epigenetic marks function via controlling the accessibility of cellular molecules to the chromatin. Some epigenetic marks such as regulatory noncoding RNA (ncRNA) can also modulate targeted genes via different mechanisms[5, 6], e.g., microRNA functions as antisense RNA degrading mRNA to negatively regulate the targets at a post-transcription level[7, 8]. This chapter will focus on the majority of the most studied epigenetic marks that control gene expression mainly through

reshaping conformation of chromatin, the essential medium for epigenetic regulation.[9] These marks are deposited, recognized, and erased by specific enzymes or proteins to acquire the inheritance and reversibility of epigenome.[4]

Chromatin is a packed format of nucleosomes which are the basic units of DNA compaction. Each nucleosome contains an octamer core of histone proteins H2A, H2B, H3, and H4 with a 147-bp DNA sequence wrapping around.[10, 11] Both nucleosome constituents carry epigenetic marks including chemical modifications and variants. For DNA, methylation of cytosine or 5-methylcytosine (5-mC) is the first and most well-studied epigenomic modification with the structure identified in 1948[12] and the function of epigenomic significance established in 1975[13, 14]. General perception of the function of DNA methylation is blocking transcription initiation and silencing gene expression when it occurs in the immediate vicinity of transcriptional starting sites (TSSs). The full scope of DNA methylation functions is more complicated as it is context-dependent.[15] DNA methylation is deposited by DNA methyltransferases, recognized by specific domain containing proteins such as methyl-CpG-binding domain (MBD) containing proteins, and removed by DNA demethylases.[16] Additionally, oxidized derivatives of 5mC including 5-hydroxymethylctosine (5-hmC)[17, 18], 5-formylcytosine (5-fC)[19], and 5-carboxylcytosine (5-caC)[19, 20] may also have regulating functions and biological consequences.[21] For histones, various modifications of the highly basic amino-terminal tails have been identified including methylation, acetylation, phosphorylation and more[22], among which methylation and acetylation on histone H3 are better explored with the former occurring mainly on lysine residues and the latter on lysine and arginine residues.[23] It has been generally identified that H3K4me3 (trimethyl-lysine 4 of H3) is associated with gene activation, H3K4me2 and H3K4me1 are enriched at enhancers or just downstream on active genes of H3K4m3, while trimethylation on

H3K9 and H3K27 are related to repressed gene expression.[24] Histone acetylation generally activates gene expression as it neutralizes the positive charges on histones, thus attenuates the otherwise strong interaction between positive-charged histones and negative charged DNA strands, relaxing the chromatin structures to be more accessible. Histone acetylation can also be recognized by specific proteins and trigger corresponding pathways.[25] Histone methyltransferases, demethylases, acetyltransferases, and deacetylases are responsible for the deposition and removal of these two marks, and specific domain containing reader proteins can differentiate these marks and bind accordingly.[16] Besides tail modifications, histone variants replace the canonical H2A and H3 and changes the properties of nucleosomes, thus also influence all chromatin-based functions including gene expression and DNA repair.[26, 27]

Not only the components of nucleosomes carry epigenomic marks, the positioning of nucleosomes is also an important epigenomic regulator of gene expression.[28] Simply speaking, nucleosomes distribute continuously and randomly within gene bodies to stabilize the genes while usually are depleted from the TSSs regions of active genes forming nucleosome-free regions (NFR) or nucleosome-depleted regions (NDR) and occupy specific locations immediately up- and downstream of these TSSs to allow the binding of transcription factors.[10, 29] Hence mapping nucleosome positioning patterns, especially their dynamic changes, reveals crucial information on chromatin function and gene regulation.[28-30]

As the understanding of these epigenomic regulators becomes further, it has now become clear that it is the complex interplay among epigenetic marks that shapes or maintains corresponding pathways and regulates nucleosome positioning and occupancy, thus alters chromatin accessibility and affects gene expression level.[9] For example, CpG-island (a stretch of DNA where the occurring frequency of CG dinucleotide sequence is higher than other regions. “p”

stands for the phosphodiester bond connecting C and G) containing promoters of active genes usually lack DNA methylation and have NFR or NDR, with surrounding nucleosomes popularly marked with H3K4me3 and commonly accompanied by the replacement of H2A with lysine-acetylated H2A.Z as such variant allows easier dissemble of or access to the nucleosome structure.[26, 31] On the other hand, DNA methylation and H3K27me3 are almost mutually exclusive silencing genes via two different mechanisms, while H3K9me plays a key role in cross talking with and maintaining DNA methylation to ensure gene stability.[9, 32-34] Therefore, it is important to overlay profiles of multiple epigenomic marks genome wide or at specific locations to fully explain their regulating mechanisms on gene expression under various cellular statuses.

Epigenomic Impact on Disease Development and Treatment. Since epigenetic marks communicate with each other to form a complex regulating network of gene expression, any genetic, environmental, and metabolic stimuli that influences single or multiple participants of epigenome may cause dysregulation of genome and in turn causes phenotype abnormality and disease development.[4] For example, p53-mutated cells (e.g., cancer cells) accumulated EZH2 (enhancer of zeste homolog 2, a histone-lysine *N*-methyltransferase) to increase H3K27me3 which induced chromatin compaction and heterochromatin formation, leading to increased resistance to DNA damage.[35] This epigenomic regulating pathway thus serves as one of the mechanisms modulating cancer cells' resistance to chemo- and radiotherapies. Additionally, tumor hypoxia reduced the activity of TET (ten-eleven translocation enzymes), leading to reduced 5-methylcytosine oxidization. This contributed to promoter hypermethylation at tumor suppressor genes, inhibiting their expression and providing a substrate for oncogenic events in solid tumors.[36] Besides cancer, epigenomic impact has also been found in other diseases such as sepsis, where the dynamic nucleosome positioning induced by the deposition and removal of positive and negative

histone marks remodeled the chromatin and controlled the induction and silencing phases of TNF-α and other inflammatory cytokines during the course of sepsis development.[30, 37]

As they are heavily involved in pathogenesis with the feature of reversibility, epigenomic marks have naturally became promising targets in diagnosis, prognosis, and therapeutics: altered or dysregulated epigenome patterns may indicate diseased statuses, and they may be reprogrammable to a healthy one.[25] Recently, CpG island methylation phenotype (CIMP) was identified as a strong risk indicator of treatment response and survival rate for patients with pediatric T-cell acute lymphoblastic leukemia (T-ALL) or relapsed BCP-ALL (B-cell precursor acute lymphoblastic leukemia). Patients with low global methylation level at the promoter CpG islands (CIMP-) had significantly inferior response to treatment with lower survival rate and higher relapse rate compared to CIMP+ patients.[38-40] Similarly, repressive mark H3K9me3 has also been pinpointed as a strong biomarker candidate and H3K9 demethylases as potential anti-cancer treatment targets. Globally enhanced H3K9 demethylase level and low H3K9me3 level was associated with melanomagenesis as such pattern allowed escape of oncogene-induced senescence. Inhibition of H3K9-active demethylases such as LSD1 and JMJD2 family members, however, rapidly restored senescence and thus controlled tumor growth.[41] Moreover, histone acetylation and corresponding removers histone deacetylases have already served as targets of histone deacetylase inhibitors (HDACi) for treatment of multiple diseases including T-cell lymphoma, multiple myeloma, and sepsis with several FDA-approved HDACi products.[25, 42] In the case of sepsis, HDACi functioned as anti-inflammatory reagent to protect the host, reduced the (over-)expression of critical immune receptors and antimicrobial products with a possible mechanism of increasing expression of the transcriptional repressor Mi-2β and enhancing the DNA-binding activity of the Mi-2/NuRD complex that repressed macrophage cytokine production.[43] As time goes by, more

epigenomic marks and regulators are being discovered, characterized, and included into the pools of potential prognostic indicators and drug targets, and related clinical development is undergoing at an extraordinary pace.[25]

Current Techniques Probing Epigenome and the Limitations. To investigate key epigenomic mechanisms impacting human health and validate epigenomic pattern changes after treatment, a variety of protocols have arisen to profile different aspects of epigenome for a thorough understanding of the embedded epigenomic landscape and infer causal relations among these epigenomic mechanisms.[44] They are coupled to the rapid growth of NGS techniques and high-throughput microarray platforms.

DNA methylation is commonly detected via chemical modification followed by sequencing. Various bisulfite sequencing (BS-seq) strategies have been widely applied to locate and quantify DNA methylation distribution along genomic DNA (**Figure 1-1**). Whole-genome bisulfite sequencing (WGBS) treats DNA with a bisulfite reagent to convert cytosine to uracil while leaving 5-mC and 5-hmC intact. Sequencing is then conducted to profile genome-wide methylation on a single nucleotide level.[45-49] Reduced representation bisulfite sequencing (RRBS) decreases cost and improves efficiency of WGBS by digesting samples with restriction enzymes (e.g. *Msp*I and *Bgl*II) to selectively enrich CpG-containing regions and extracting the fragmented DNA with appropriate sizes (<600 bp) prior to bisulfite treatment.[50-52] WGBS and RRBS cannot differentiate 5-mC from 5-hmC, though. Tet-assisted bisulfite sequencing (TAB-seq) specifically detects 5-hmC at single-base resolution, determining its abundance at each modification site by protecting hydroxymethyl- group with glycosylation and oxidizing other cytosine residues with Tet enzyme, the result of which leaves only 5hmC unaffected upon bisulfite treatment.[19, 53, 54] Another modified BS strategy, oxidative bisulfite sequencing (oxBS-Seq), is also capable of

differentiating between 5-mC and 5-hmC by specifically oxidizing 5-hmC before bisulfite treatment to obtain a positive readout of pure 5-mC. 5-hmC readout is inferred by comparing oxBS-seq with parallel-run regular BS-seq data,.[55, 56] Compared to TAB-seq, oxBS-seq does not require expensive enzymes and avoids the potential inefficiency in glycosylation of 5-hmC or enzymatic oxidization of 5mC.[48, 56] However, oxBS-seq requires two runs of sequencing thus significantly increases the financial and data processing burden. Beyond techniques targeting 5-mC and 5-hmC, strategies detecting 5-fC and 5-caC have also emerged, such as 5-fC chemically assisted bisulphite sequencing (fCAB-seq) for 5-fC mapping and chemical modification-assisted bisulphite sequencing (CAB-seq) for 5-caC mapping.[57] As a consequence of their relative newness, these techniques have limited established results compared to those for 5-mC and 5-hmC.

With the applications of these detection methods, the mechanisms of DNA methylation regulating gene expression have been further understood. WGBS and RRBS have been used to reveal dynamic methylome changes upon effects from non-shared or unique environmental factors,[58] map genome-wide DNA methylation in mammalian ovaries,[59] and track genome-wide DNA methylation reprogramming and resulted biological consequences during mouse primordial germ cell expansion and migration.[60] TAB-seq and oxBS-seq have been used to understand the biological functions of 5-hmC in cellular memory reprogramming, cancer development, and other diseases[61-64]. Moreover, BS-seq based strategies have been undergoing modification and optimization to enable higher coverage and accuracy in mapping DNA methylome.

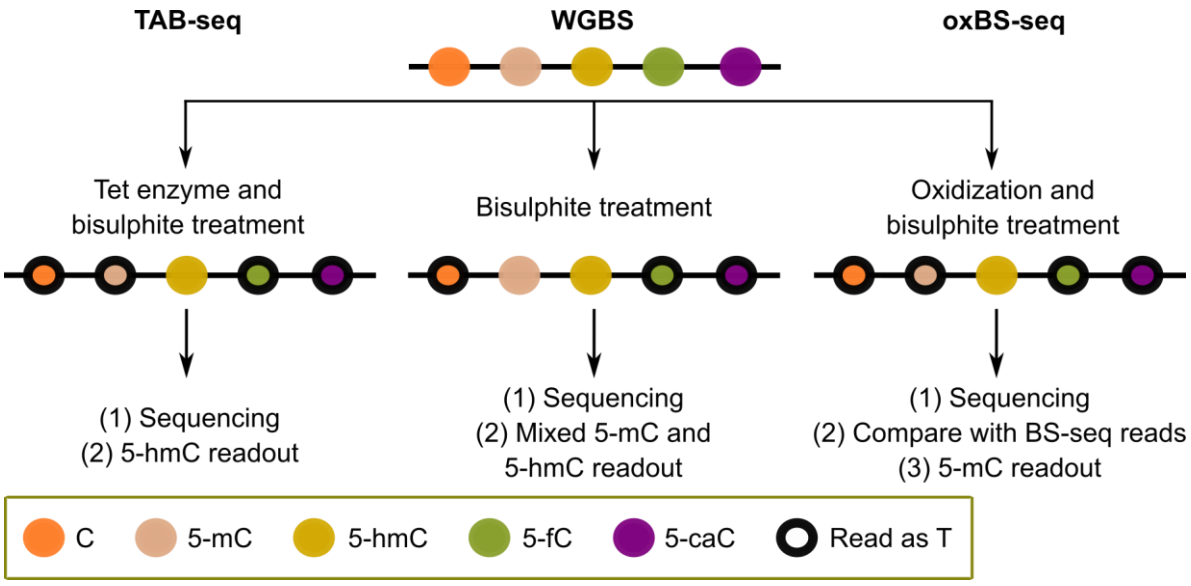

Figure 1-1 Selected protocols of chemical modification-based DNA methylation detections.

Histone modifications as well as DNA methylation can be detected through affinity-based methods followed by sequencing. In addition to chemical-based methods, affinity-based protocols using antibodies to specifically recognize modifications of interest on DNA-associated histone proteins and DNA methylation have been developed and widely used in laboratory epigenetic research (**Figure 1-2**).

Methylated DNA immunoprecipitation (MeDIP) was developed to enrich 5-mC DNA regions with the highly efficient binding from 5-mC directed antibodies. The enriched product is then sequenced, leading to a powerful MeDIP-seq strategy that generates high-resolution whole-genome DNA methylation profiles.[65, 66] Similarly, antibodies targeting 5-hmC, 5-fC, and 5-caC have also been utilized to locate respective modified cytosine residues, providing multiple variants of DIP-seq strategies (i.e., hMeDIP-seq, 5-fC-DIP-seq, and 5-caC-DIP-seq).[67-69] Methods using methyl-CpG binding domain (MBD) of methyl-CpG binding domain protein 2 (MBD2) to precipitate methylated DNA regions have also been applied to identify DNA methylation pattern integrated with sequencing (i.e. MBD-seq based methods).[70-72] Critical roles of DNA methylation

and its variations in regulating chromatin accessibility, altering gene expression and affecting disease/caner development have been clarified by applying one or multiple of these methods to reveal their distribution and changes along the reference genome sequences.[63, 73-75]

Affinity-based immunoprecipitation was also applied to analyze chemical modifications (e.g., methylation and acetylation) as well as binding of transcription factors and chromatic remodeling complexes to histones that influence transcriptional regulation. The gold standard for such protein-DNA interaction is chromatin immunoprecipitation followed by sequencing (ChIP-seq) (**Figure 1-2**, right protocol analyzed with sequencing). ChIP-seq generally involves (1) optional crosslinking of DNA and proteins to preserve their dynamic interaction if not strong, (2) shearing of chromatin into mono- and di-nucleosome sizes (about 150-300 bp) by sonication and/or micrococcal nuclease, (3) immunocapture of the modification of interest with specific antibodies, and (4) purification of captured DNA for library preparation and sequencing.[76-78] The quality of the sequenced results can be assessed by comparing to a control library obtained with nonspecific antibody following the same protocol. Due to the size heterogeneity resulting from inconsistencies in DNA shearing and potential contamination from nonspecific-bound DNA, etc., conventional ChIP-seq suffers from low resolution in target mapping, and may not provide a complete set of protein binding locations.[48] Additionally, a lack of normalization protocols that would allow for quantitative comparison between samples is also a shortcoming of this powerful method.[79, 80]

To address these issues, variations of ChIP-seq have been constructed. ChIP-exo utilizes exonuclease treatment to narrow the region of protein binding on ChIP DNA sequences, enhancing mapping resolution of bound locations to the single-nucleotide level.[79] ChIP with reference exogenous genome (ChIP-Rx) allows normalization of sequencing reads across cell populations

by adding a defined quantity of a reference epigenome on a per-cell basis, thus enabling quantitative comparison among multiple ChIP-seq runs.[80] Although they are utilized on macroscale of samples commonly and still contain space for further optimization, ChIP-seq and derivative methods allow for identification and validation of modifiers and regulators related to aging, cellular development, disease progression, and other crucial biological processes, along with the disclosure of underlying epigenetic regulating mechanisms.[81-86]

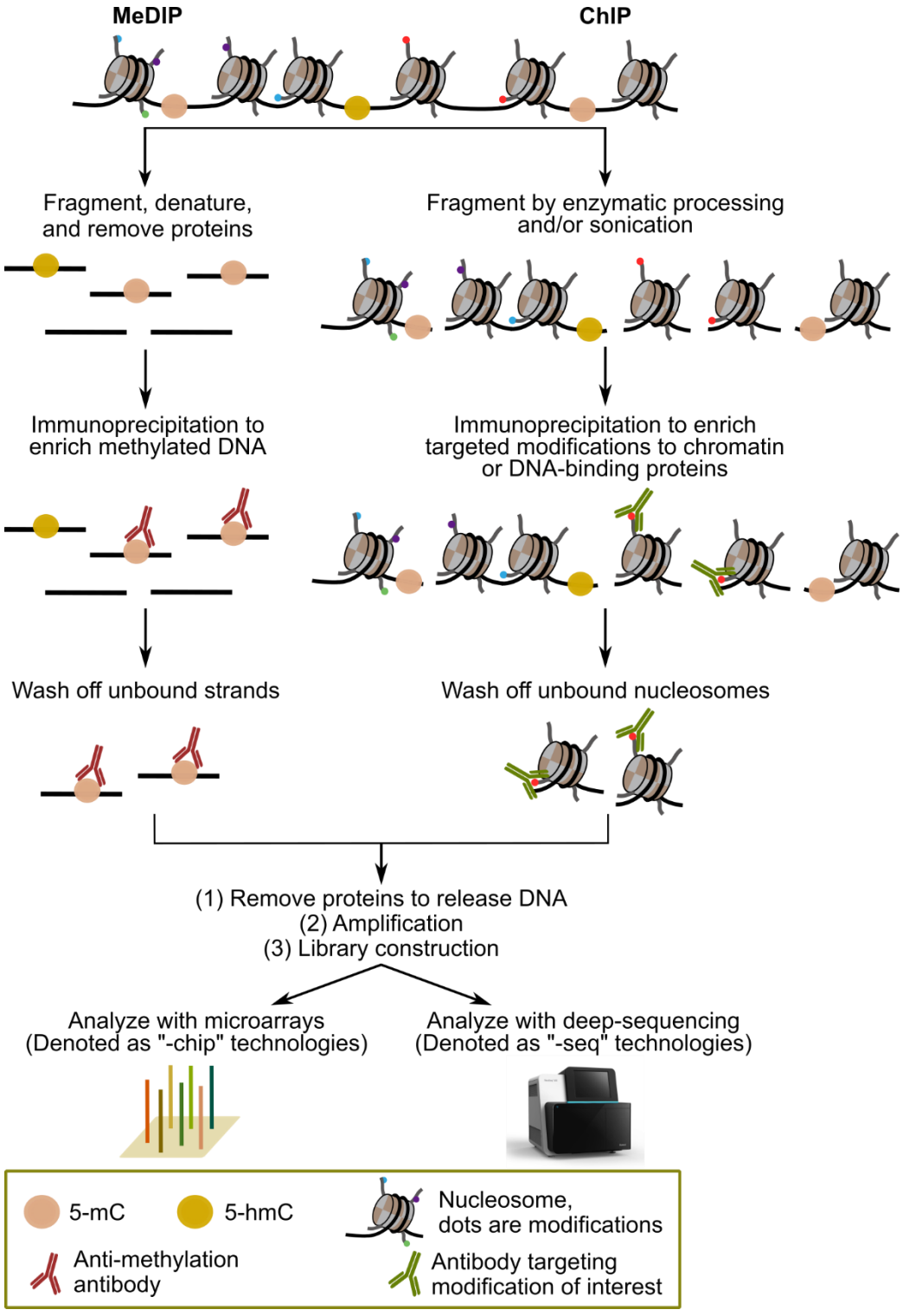

Figure 1-2 Selected protocols of affinity-based chromatin modification detections.

Nucleosome distribution and open chromatin can be assessed by enzymatic and chemical processing followed by sequencing. Various methods have been developed to map these important features exploiting different enzymes and chemicals to expose nucleosomal or regulatory regions of interest (**Figure 1**-**3**).

Micrococcal nuclease (MNase) is an endo-exonuclease that has long been used to digest chromatin with minimal sequence preference into (mono-)nucleosomes as it cleaves exposed DNA ends until encountering a barrier (e.g., nucleosome).[87, 88] Combining MNase digestion with high-throughput sequencing (MNase-seq) enables localization of DNA-binding proteins (e.g., transcriptional factors) and nucleosome positioning at superior resolution as high as single base pair.[89] DNase-seq, on the other hand, exposes non-occluded accessible regions on genomic sequences by limited digestion of the endonuclease DNase I, characterizing the opposite regions on the genome when compared to MNase-seq.[90] This allows the mapping of genetic *cis*-regulatory elements (e.g., promoters, enhancers, and silencers) in open chromatin structures, which reveals dynamic changes of acquisition and loss of transcription competence of genes.[91, 92] Both MNase-seq and DNase-seq separate enzymatic processing of chromatin from adaptor ligation to purified DNA. Recently, an assay for transposase-accessible chromatin using sequencing (ATAC-seq) has been developed, utilizing hyperactive Tn5 transposase to fragment chromatin and tag DNA with sequencing-compatible adaptors simultaneously.[93] ATAC-seq enables fast and potentially simultaneous genome-wide mapping of active regulatory elements, nucleosome positioning, and chromatin accessibility, thus has gained growing popularity.[94]

Apart from enzymes, chemicals can also be used to process chromatin. When treating chromatin with formaldehyde, a common crosslinking reagent, nucleosome-depleted regions obtain lower efficiency as fewer proteins are available to be chemically bonded with DNA

sequences comparing to nucleosome-occupied sites. Therefore, nucleosome-depleted DNA is enriched to aqueous phase after phenol-chloroform extraction. This leads to a strategy named formaldehyde-assisted isolation of regulatory elements (FAIRE).[95] FAIRE-seq provides insight on active regulatory elements abundant in open chromatin, serving as another alternative to DNase-seq and ATAC-seq without the necessity of relying on any enzymatic digestion.[96]

When applied individually and integratedly to studies, these methods have enhanced the community's understanding in the dynamics of and relation between nucleosome occupancy and chromatin remodeling during various biological processes and disease development, including human hematopoiesis, leukemia development, and oncogenic state forming.[86, 97-99]

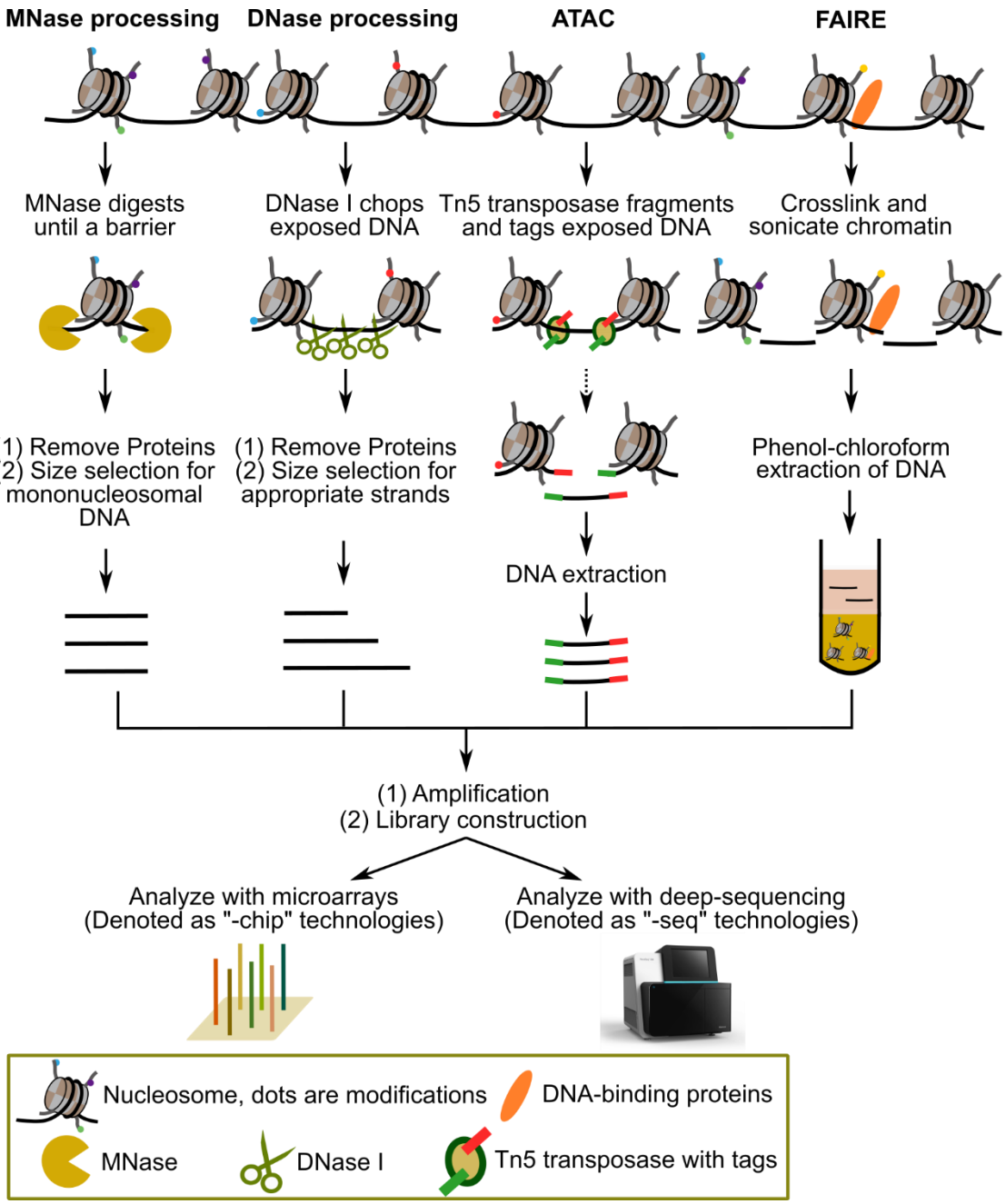

Figure 1-3 Selected protocols to analyze nucleosome distribution and open chromatin.

Epigenomic marks can also be evaluated by hybridization-based technologies. Sequencing-based methods have revolutionized our understanding of epigenetic regulation on a genome-wide level. However, translating these protocols to large clinical cohorts can prove challenging, where cost, throughput, and reproducibility all become necessary considerations. As

a result, hybridization-based oligonucleotide microarray method has been developed for more cost-effective detections.[52] Microarray methods have grasped wide popularity in detecting DNA methylation and histone modifications, where they have been used to assess the progression and effects of adiposity[100], quantify metabolism's contribution to tumor development[101], mapping DNA methylome footprint during B cell differentiation[102] and lung cancer[103], and more. This popularity is reflected in the fact that a variety of commercial kits are available to researchers. The representative product is the Infinium Methylation Assays from illumina. Infinium HumanMethylation450 BeadChip Kit (referred to as 450 kit below, recently replaced by Infinium Methylation EPIC Kit) provides more than 485,000 methylation cites per samples at single-nucleotide resolution across the genome. Upon bisulfite treatment, the assay interrogates the chemically modified sites with two probes differentiating methylated and unmethylated loci, allowing analysis of up to 96 samples simultaneously. Tiling microarray assays that combine immunoprecipitation or chemical modification with hybridization include ChIP-chip, MeDIP-chip, and DNase-chip ("chip" indicates microarray in the denomination, also seen in **Figure 1-2 and 3**). After targeting and isolating DNA associated with modification of interest via ChIP, bisulfite treatment, or other chemical modification methods as mentioned above, purified DNA hybridizes with complimentary probes to be identified instead of being sequenced. Applications of these arrays are commonly found in a host of studies including identifying specialized chromatin domains in transcriptional regulation[104], recognizing protein markers for psychiatric disorder development[105], and marking cell-specific epigenomic sites.[106]

Translating Lab Technologies to Clinical Settings. Among all the epigenomic markers that have been investigated with the various methods introduced above, DNA methylation is the most commonly studied epigenetic marker in mammalian genomes.[107] Commercially available assays

aimed to lower the experimental burden through increased automation and reduced manual processing have been developed to study DNA methylation in clinical settings.[108, 109] Hence, the focus of the discussion below will outline research efforts to date aiming at automating the analysis of protein-DNA interactions with the goal of ultimately translating these technologies into a clinical setting.

Limitations of ChIP-seq and Variants. While ChIP-seq and its variants have traditionally been the work horse for protein-nucleic acid interactions, these protocols have limitations that prevent their use in clinical settings besides those discussed above when introducing ChIP-seq and its derivatives. One major challenge is the large input sample size requirement. Traditionally, ChIP-seq requires more than 10^6 cells as starting materials for one mark to ensure effective enrichment of related DNA, which is unfeasible for small sizes of clinical samples or subpopulations of rare cells.[110] A second challenge is the lack of sensitivity to cell heterogeneity. Epigenome is dynamic, tissue specific, and even single-cell specific especially in situations where cell heterogeneity is common (e.g., tumors).[111, 112] Due to the aggregation of DNA from all cells present in the sample of interest, conventional ChIP buries possible signals from small numbers of cells carrying abnormal or novel features.[113, 114] Third, robustness and reproducibility of ChIP-seq need to be enhanced. ChIP-seq experiments require high-quality chromatin, validated antibodies, carefully selected commercial kits for ChIP DNA purification, and the appropriate tubes for ChIP DNA preservation.[115] They also rely largely on an operator's skill to complete a large number of manual processing steps making the protocols lengthy and cumbersome. Fourth, ChIP-seq experiments are inherently low throughput, where they target only one protein or modification per sample volume. When combined, these factors complicate the use of conventional ChIP-seq in

routine clinical practices that study large patient cohorts such as monitoring both healthy and abnormal samples across human life span.[52]

Efforts have been made to improve conventional ChIP-seq by reducing cell numbers, shortening DNA preparation time, and improving easeness/reducing sample loss by doing as many steps in a same tube as possible or carrying out the procedure in 96-well plates.[116-119] However, none of these examples reduced the large numbers of steps needed or automated the entire ChIP process, nor did they allowed single-cell studies.

Promising Approaches of Automated Epigenomic Analysis. Recent research efforts have turned to microfluidic tools as powerful alternatives to macroscale methods when it comes to develop an automated ChIP-seq platform. It is now widely known that microfluidics offers miniaturization of reagent volume, capability of parallelization and integration, and potential of automation, which are all particularly well-suited to sample-constrained, labor-intensive, and operator-dependent epigenetic analyses.[120] Additionally, because of the microscale sizes, microfluidics facilitates single cell analysis, which will benefit investigation of cell epigenomic heterogeneity that has mostly been averaged out when using macroscale methods. The toolbox of single cell manipulation and processing on microfluidics has been expanding in the past few years with the emergence of fantastic proof-of-concept work including highly efficient single cell encapsulation in microfluidic droplets (**Figure 1-4a**)[121], high-throughput single cell reverse transcription PCR (RT-PCR) (**Figure 1-4b**)[122-124], and single-cell sequencing with comparable read count and improved sensitivity compared to tube-based protocols (**Figure 1-4c**).[125-127] By adapting the relevant principles, many exciting studies have been completed to facilitate microfluidic epigenomic analysis.

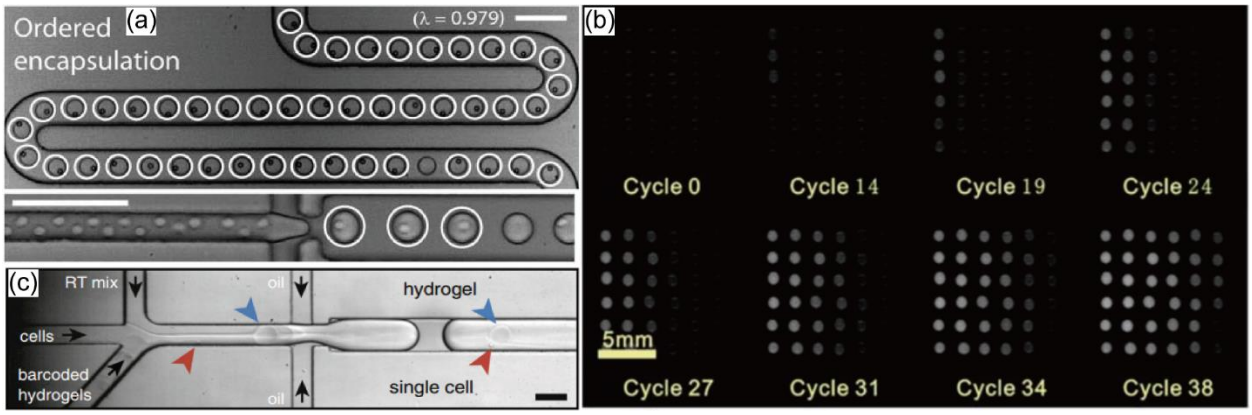

Figure 1-4 Selected microfluidic tools for epigenomic analysis. (a) Ordered single-bead (top) and single-cell encapsulation (bottom).[121] Scale bars: 100 μm. Adapted from Ref 121 with permission of The Royal Society of Chemistry. (b) Single cell droplet array RT-PCR. Fluorescence images of droplet array at different thermal cycles. The input of mir-122 in each 6×6 array was 9.6×108 (left lane1, top three droplets), 9.6×107 (left lane1, bottom three droplets), 9.6×106 (lane 2), 9.6×105 (lane 3), 9.6×104 (lane 4), 9.6×103 (lane 5), 9.6×102 (lane 6, top three droplets) and 0 copies per droplet (lane 6, bottom three droplets), respectively.[122] Adapted from Ref 122 with permission of The Royal Society of Chemistry. (c) Cells and reagent encapsulation in DNA barcoding-single cell whole transcriptome sequencing with droplets.[126] Scale bar: 100 μm. Adapted from Ref 126 with permission of Elsevier.

The Quake group has been piloting in enabling microfluidic low-input cells and single-cell epigenomic sequencing and profiling. They have developed automated microfluidic device-based ChIP protocol (AutoChIP) that allows ChIP DNA enrichment from 2000 cells compared to conventional methods.[128, 129] Multi-layered and valve-assisted ring structures were designed to control sample loading, bead washing, and DNA eluting. By integrating multiple AutoChIP structures, they have also achieved a high throughput, automated microfluidic device for ChIP (HTChIP) capable of processing 10,000-cell equivalent chromatin samples in each parallel structure, running 14 ChIP and 2 controls simultaneously (**Figure 1-5a**).[130]

The Lu lab has also been developing epigenomic assays for fewer cells by exploiting microfluidic chamber structures. By using valves to control chromatin loading and capture DNA elution, they have shown that fragmented chromatin from 100 cells can be treated microfluidically

with high enough quality for subsequent DNA analysis (**Figure 1-5b**).[131] Further integration of sonication and immunoprecipitation enabled on-chip shearing of chromatin in addition to DNA capture (**Figure 1-5c**).[132] Later, they further modified the original design and obtained a fluidized-bed-based device for parallel analyses on two histone marks with the same batch of sheared chromatin.[133]

These promising examples show the potential of microfluidics to facilitate automated, low-input and single-cell epigenomic studies. However, they have their own limitations. The microfluidic chamber-based devices such as those from the Quake lab and the Lu lab, though allowed low-input samples, limited the sample size or antibody-functionalized beads number by an intrinsic factor: the device size. Only the specific volume of cell suspension or fragmented chromatin that fits the chamber can be processed, otherwise part of the samples had to be proactively expelled causing inevitable sample loss; only the amount of chromatin that is optimized against the number of beads loaded into the device can be processed, otherwise the immunocapture efficiency would be jeopardized. Thus, such design limits the sample size to the smaller end and has low flexibility/tolerance of sample size variance. Moreover, fabrication and operation of these multi-valved devices is complicated and thus not user-friendly. To improve on these aspects, another format of microfluidics, droplet microfluidics, may represents a better option, as droplets can be continuously generated allowing flexible sample sizes or magnetic bead numbers. Indeed, the Weitz lab and colleagues have applied droplet microfluidics to achieve automated single-cell ChIP-seq (Drop-ChIP) that can handle samples with indeterminate sizes and process thousands of cells individually within minutes. Starting with single cell encapsulation followed by simultaneous introduction of chromatin fragmentation and DNA immunocapture reagents to cell-containing droplets, Drop-ChIP used DNA barcoding to identify different cells

and analyze a mixture of three different cell types at single-cell resolution using deconvoluted chromatin state mapping, leading to cell subpopulation elucidation (**Figure 1-5d**).[134] The specificity and information content accuracy were acceptable, as ~50% of reads could be aligned to known positive sites, and coverage of aggregated reads from 50 cells was comparable to conventional profiles. On the other hand, the coverage per cell was sparse on the order of 1000 unique reads. Additionally, this device only automated cell encapsulation with lysis/digestion reagent and DNA barcode indexing; the two key steps of immunoprecipitation and ChIP DNA extraction were still carried out manually off the device. Furthermore, potential repetitions in DNA barcoding introduced possible false identification of cell types, and thus could confound the results.

To summarize, droplet microfluidics is a promising technique for flexible sample sized epigenomic assays especially for low-input to single-cell samples. With the only available example of ChIP in droplets containing areas for improving, it will be exciting to apply this technique to the methodology development for various epigenomic assays with the motivation to automate them for practical applications in clinical settings and promote epigenomic diagnosis, prognosis, and therapeutics. A long-term goal is realizing droplet microfluidic-based, automated epigenome profiling and screening for every single patient, realizing precision medicine from the aspect of epigenomics.[110, 135]

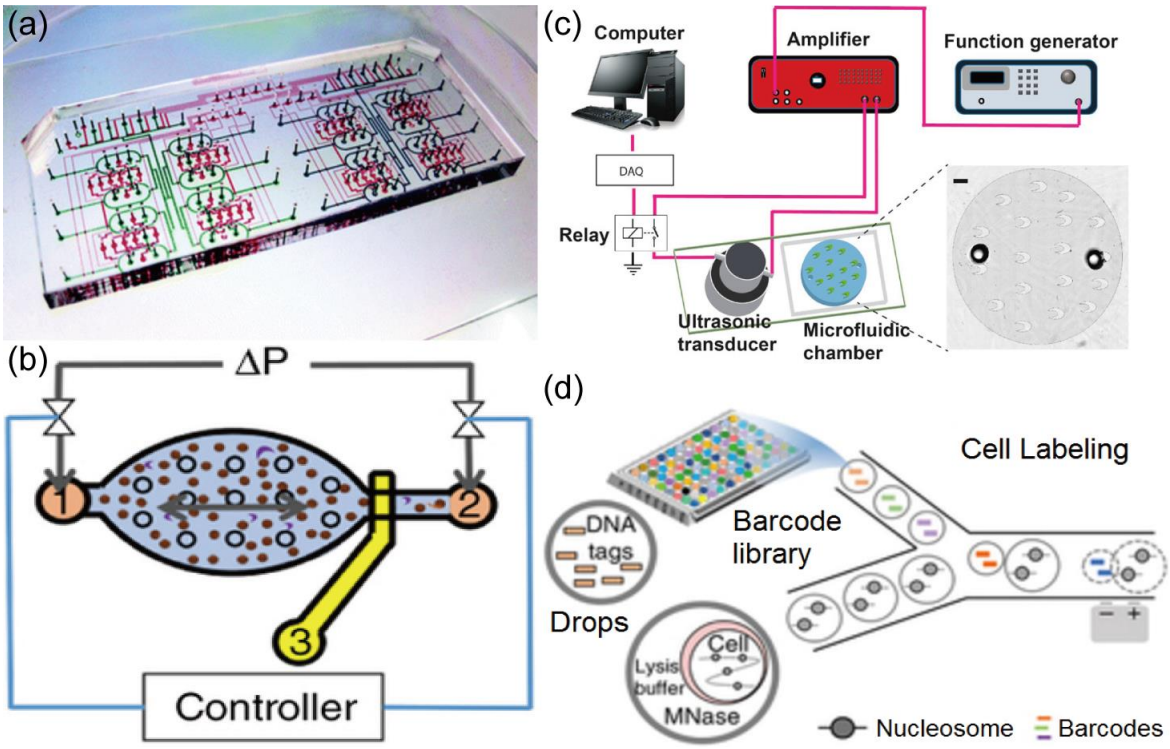

Figure 1-5 Selected microfluidic devices for epigenomic analysis. (a) A ring-structured platform for performing high-throughput ChIP (HTChIP) screening measurements of 16 different targets simultaneously.[130] Adapted from Ref 130 with permission of The Royal Society of Chemistry. (b) A microfluidic oscillatory washing–based ChIP-seq (MOWChIP-seq) device for epigenetic profiling with 100 cells.[131] Adapted from Ref 131 with permission of Springer Nature. (c) An integrated device for on-chip sonication and immunoprecipitation.[132] Adapted from Ref 132 with permission of The American Chemical Society. Further permissions should be directed to the ACS. (d) A droplet-based microfluidic device for single-cell ChIP-seq (Drop-ChIP) integrating cell encapsulation, reagents addition, and DNA barcoding to analyze thousands of cells at single-cell resolution.[134] Adapted from Ref 134 with permission of Springer Nature.

Droplet Microfluidics

Droplet microfluidics is a rapidly developing interdisciplinary field of research combining physics, chemistry, biology, and microsystems engineering. It is becoming a valuable tool for various applications such as single-cell analysis, complex biological and chemical assays, diagnostics, DNA sequencing, and drug screening.[136-140]

The basic principle to generate droplets is segmenting aqueous streams with a continuous, immiscible, and inert oil flow, and each droplet functions as an independent microreactor.[141, 142]

The volume of the droplets typically range from femtoliters to nanoliters, leading to extremely small reagent consumption suitable for rare samples such as stem cells and circulating cancer cells.[143-149] Droplet sizes are tunable by changing flow rate values and ratios between continuous and dispersed phases.[150] Droplet microfluidics also contains single cell analytical capability as individual cells can be encapsulated and processed in single droplets.[121, 134, 151-156]

Reliable and precise manipulation of individual droplet has emerged recently covering droplet coalescence, splitting, injection, and content mixing.[157-168] Additionally, previous studies have shown that mixing occurs very rapidly within the droplets due to convective flow profile inside and when assisted by serpentine design, allowing for more efficient chemical and biological reactions and thus shorter operation time for experiments.[141, 169-171] This is highly advantageous over conventional continuous-flow microfluidics, which offers low Reynolds numbers and laminar flows, resulting in rather slower mixing predominated by molecular diffusion and would need additional mechanism for mixing.[141]

With all the favorable features discussed above, it is feasible to utilize droplet microfluidics to automate the entire workflows of epigenomic assays with size-flexible and low-input samples to approach the expectations in clinical applications. The various manipulations of droplets offer abundant options to miniaturize the massive manual processing involved in these assays. The potential for single cell analysis also opens the door for future epigenomic heterogeneity studies on size-limited samples.

Dissertation Overview

In this thesis, the current status of epigenetics and the motivation of using droplet microfluidics to automate epigenomic assays were discussed in Chapter 1. The devices used in this

thesis were designed, fabricated, and characterized with the detailed protocol described in Chapter 2. Chapter 3 to 5 described the automation of sample preparation for three epigenomic assays (MNase-seq, MBA/MNase-seq dual assays, and ChIP-seq) using droplet microfluidics because of its favorable features including isolation of samples for reduced nonspecific binding, rapid mixing, sample scalability, and the potential for single cell studies. Overall, this thesis provided the beginning of automating various epigenomic assays in droplet microfluidics for practical applications. Preliminary results from clinical samples were discussed in Chapter 6, providing compelling evidence for the applicability of droplet microfluidic platforms in clinical epigenomic studies. Future work was pointed out in Chapter 6 to further the accomplishments achieved in this thesis.

References

1. C. H. Waddington, *Endeavour*, 1942, **1**, 18-20.
2. E. Jablonka and E. Lamm, *International journal of epidemiology*, 2012, **41**, 16-20.
3. C. Dupont, D. R. Armant and C. A. Brenner, *Seminars in reproductive medicine*, 2009, **27**, 351-357.
4. W. A. Flavahan, E. Gaskell and B. E. Bernstein, *Science*, 2017, **357**.
5. J. Chen, S. Liu and X. Hu, *Cell death discovery*, 2018, **4**, 50.
6. D. Holoch and D. Moazed, *Nature reviews. Genetics*, 2015, **16**, 71-84.
7. H. Guo, N. T. Ingolia, J. S. Weissman and D. P. Bartel, *Nature*, 2010, **466**, 835-840.
8. H. Wang, R. Peng, J. Wang, Z. Qin and L. Xue, *Clinical epigenetics*, 2018, **10**, 59.
9. Y. Atlasi and H. G. Stunnenberg, *Nature reviews. Genetics*, 2017, **18**, 643-658.
10. C. Jiang and B. F. Pugh, *Nature reviews. Genetics*, 2009, **10**, 161-172.
11. V. W. Zhou, A. Goren and B. E. Bernstein, *Nature reviews. Genetics*, 2011, **12**, 7-18.
12. R. D. Hotchkiss, *The Journal of biological chemistry*, 1948, **175**, 315-332.
13. R. Holliday and J. E. Pugh, *Science*, 1975, **187**, 226-232.
14. A. D. Riggs, *Cytogenetics and cell genetics*, 1975, **14**, 9-25.
15. Y. Yin, E. Morgunova, A. Jolma, E. Kaasinen, B. Sahu, S. Khund-Sayeed, P. K. Das, T. Kivioja, K. Dave, F. Zhong, K. R. Nitta, M. Taipale, A. Popov, P. A. Ginno, S. Domcke, J. Yan, D. Schubeler, C. Vinson and J. Taipale, *Science*, 2017, **356**.
16. Z. Chen, S. Li, S. Subramaniam, J. Y. Shyy and S. Chien, *Annu Rev Biomed Eng*, 2017, **19**, 195-219.
17. S. Kriaucionis and N. Heintz, *Science*, 2009, **324**, 929-930.

18. M. Tahiliani, K. P. Koh, Y. Shen, W. A. Pastor, H. Bandukwala, Y. Brudno, S. Agarwal, L. M. Iyer, D. R. Liu, L. Aravind and A. Rao, *Science*, 2009, **324**, 930-935.
19. S. Ito, L. Shen, Q. Dai, S. C. Wu, L. B. Collins, J. A. Swenberg, C. He and Y. Zhang, *Science*, 2011, **333**, 1300-1303.
20. P. Brazauskas and S. Kriaucionis, *Nat Chem*, 2014, **6**, 1031-1033.
21. A. Klungland and A. B. Robertson, *Free radical biology & medicine*, 2017, **107**, 62-68.
22. A. J. Bannister and T. Kouzarides, *Cell Res*, 2011, **21**, 381-395.
23. E. E. Yörüker, S. Holdenrieder and U. Gezer, *Translational Cancer Research*, 2018, **7**, S185-S191.
24. H. Kimura, *Journal of human genetics*, 2013, **58**, 439-445.
25. A. Ganesan, *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 2018, **373**.
26. D. Corujo and M. Buschbeck, *Cancers*, 2018, **10**.
27. D. Filipescu, S. Müller and G. Almouzni, *Annual Review of Cell and Developmental Biology*, 2014, **30**, 615-646.
28. W. K. M. Lai and B. F. Pugh, *Nature reviews. Molecular cell biology*, 2017, **18**, 548-562.
29. K. Maehara and Y. Ohkawa, *Scientific reports*, 2016, **6**, 19620.
30. M. El Gazzar, T. Liu, B. K. Yoza and C. E. McCall, *The Journal of biological chemistry*, 2010, **285**, 1259-1271.
31. P. A. Jones, *Nature reviews. Genetics*, 2012, **13**, 484-492.
32. S. B. Rothbart, K. Krajewski, N. Nady, W. Tempel, S. Xue, A. I. Badeaux, D. Barsyte-Lovejoy, J. Y. Martinez, M. T. Bedford, S. M. Fuchs, C. H. Arrowsmith and B. D. Strahl, *Nature structural & molecular biology*, 2012, **19**, 1155-1160.
33. X. Liu, Q. Gao, P. Li, Q. Zhao, J. Zhang, J. Li, H. Koseki and J. Wong, *Nature communications*, 2013, **4**, 1563.
34. X. Yang, B. Hu, Y. Hou, Y. Qiao, R. Wang, Y. Chen, Y. Qian, S. Feng, J. Chen, C. Liu, G. Peng, F. Tang and N. Jing, *Cell Res*, 2018, **28**, 593-596.
35. G. Kuser-Abali, L. Gong, J. Yan, Q. Liu, W. Zeng, A. Williamson, C. B. Lim, M. E. Molloy, J. B. Little, L. Huang and Z.-M. Yuan, *Proceedings of the National Academy of Sciences*, 2018, **115**, 3452-3457.
36. B. Thienpont, J. Steinbacher, H. Zhao, F. D'Anna, A. Kuchnio, A. Ploumakis, B. Ghesquiere, L. Van Dyck, B. Boeckx, L. Schoonjans, E. Hermans, F. Amant, V. N. Kristensen, K. Peng Koh, M. Mazzone, M. Coleman, T. Carell, P. Carmeliet and D. Lambrechts, *Nature*, 2016, **537**, 63-68.
37. S. H. Park, K. Kang, E. Giannopoulou, Y. Qiao, G. Kim, K. H. Park-Min and L. B. Ivashkiv, *Nature immunology*, 2017, **18**, 1104-1116.
38. M. Borssen, L. Palmqvist, K. Karrman, J. Abrahamsson, M. Behrendtz, J. Heldrup, E. Forestier, G. Roos and S. Degerman, *PloS one*, 2013, **8**, e65373.
39. M. Borssen, Z. Haider, M. Landfors, U. Noren-Nystrom, K. Schmiegelow, A. E. Asberg, J. Kanerva, H. O. Madsen, H. Marquart, M. Heyman, M. Hultdin, G. Roos, E. Forestier and S. Degerman, *Pediatric blood & cancer*, 2016, **63**, 1185-1192.
40. M. Borssen, J. Nordlund, Z. Haider, M. Landfors, P. Larsson, J. Kanerva, K. Schmiegelow, T. Flaegstad, O. G. Jonsson, B. M. Frost, J. Palle, E. Forestier, M. Heyman, M. Hultdin, G. Lonnerholm and S. Degerman, *Clinical epigenetics*, 2018, **10**, 31.
41. Y. Yu, K. Schleich, B. Yue, S. Ji, P. Lohneis, K. Kemper, M. R. Silvis, N. Qutob, E. van Rooijen, M. Werner-Klein, L. Li, D. Dhawan, S. Meierjohann, M. Reimann, A. Elkahloun,

S. Treitschke, B. Dorken, C. Speck, F. A. Mallette, L. I. Zon, S. L. Holmen, D. S. Peeper, Y. Samuels, C. A. Schmitt and S. Lee, *Cancer cell*, 2018, **33**, 322-336 e328.
42. E. Ciarlo, A. Savva and T. Roger, *International journal of antimicrobial agents*, 2013, **42 Suppl**, S8-12.
43. T. Roger, J. Lugrin, D. Le Roy, G. Goy, M. Mombelli, T. Koessler, X. C. Ding, A. L. Chanson, M. K. Reymond, I. Miconnet, J. Schrenzel, P. Francois and T. Calandra, *Blood*, 2011, **117**, 1205-1217.
44. S. H. Stricker, A. Koferle and S. Beck, *Nature reviews. Genetics*, 2017, **18**, 51-66.
45. R. Shapiro, R. E. Servis and M. Welcher, *Journal of the American Chemical Society*, 1970, **92**, 422-424.
46. M. J. Ziller, K. D. Hansen, A. Meissner and M. J. Aryee, *Nature methods*, 2015, **12**, 230-232, 231 p following 232.
47. H. Hayatsu, Y. Wataya, K. Kai and S. Iida, *Biochemistry*, 1970, **9**, 2858-2865.
48. J. Shin, G. L. Ming and H. Song, *Nature neuroscience*, 2014, **17**, 1463-1475.
49. M. Frommer, L. E. McDonald, D. S. Millar, C. M. Collis, F. Watt, G. W. Grigg, P. L. Molloy and C. L. Paul, *Proceedings of the National Academy of Sciences*, 1992, **89**, 1827-1831.
50. A. Meissner, A. Gnirke, G. W. Bell, B. Ramsahoye, E. S. Lander and R. Jaenisch, *Nucleic acids research*, 2005, **33**, 5868-5877.
51. H. Gu, Z. D. Smith, C. Bock, P. Boyle, A. Gnirke and A. Meissner, *Nature protocols*, 2011, **6**, 468-481.
52. H. Yan, S. Tian, S. L. Slager, Z. Sun and T. Ordog, *American journal of epidemiology*, 2016, **183**, 96-109.
53. M. Yu, G. C. Hon, K. E. Szulwach, C. X. Song, P. Jin, B. Ren and C. He, *Nature protocols*, 2012, **7**, 2159-2170.
54. M. Yu, G. C. Hon, K. E. Szulwach, C. X. Song, L. Zhang, A. Kim, X. Li, Q. Dai, Y. Shen, B. Park, J. H. Min, P. Jin, B. Ren and C. He, *Cell*, 2012, **149**, 1368-1380.
55. M. J. Booth, M. R. Branco, G. Ficz, D. Oxley, F. Krueger, W. Reik and S. Balasubramanian, *Science*, 2012, **336**, 934-937.
56. M. J. Booth, T. W. Ost, D. Beraldi, N. M. Bell, M. R. Branco, W. Reik and S. Balasubramanian, *Nature protocols*, 2013, **8**, 1841-1851.
57. N. Plongthongkum, D. H. Diep and K. Zhang, *Nature reviews. Genetics*, 2014, **15**, 647-661.
58. S. Busche, X. Shao, M. Caron, T. Kwan, F. Allum, W. A. Cheung, B. Ge, S. Westfall, M. M. Simon, A. Barrett, J. T. Bell, M. I. McCarthy, P. Deloukas, M. Blanchette, G. Bourque, T. D. Spector, M. Lathrop, T. Pastinen and E. Grundberg, *Genome biology*, 2015, **16**, 290.
59. X. L. Yuan, N. Gao, Y. Xing, H. B. Zhang, A. L. Zhang, J. Liu, J. L. He, Y. Xu, W. M. Lin, Z. M. Chen, H. Zhang, Z. Zhang and J. Q. Li, *Scientific reports*, 2016, **6**, 22138.
60. S. Seisenberger, S. Andrews, F. Krueger, J. Arand, J. Walter, F. Santos, C. Popp, B. Thienpont, W. Dean and W. Reik, *Molecular cell*, 2012, **48**, 849-862.
61. P. Yu, L. Ji, K. J. Lee, M. Yu, C. He, S. Ambati, E. C. McKinney, C. Jackson, C. A. Baile, R. J. Schmitz and R. B. Meagher, *PloS one*, 2016, **11**, e0154949.
62. C. J. Mariani, J. Madzo, E. L. Moen, A. Yesilkanal and L. A. Godley, *Cancers*, 2013, **5**, 786-814.
63. A. Zhubi, Y. Chen, E. Dong, E. H. Cook, A. Guidotti and D. R. Grayson, *Translational psychiatry*, 2014, **4**, e349.

64. X. Li, Y. Liu, T. Salz, K. D. Hansen and A. Feinberg, *Genome research*, 2016, **26**, 1730-1741.
65. M. Weber, J. J. Davies, D. Wittig, E. J. Oakeley, M. Haase, W. L. Lam and D. Schubeler, *Nature genetics*, 2005, **37**, 853-862.
66. T. A. Down, V. K. Rakyan, D. J. Turner, P. Flicek, H. Li, E. Kulesha, S. Graf, N. Johnson, J. Herrero, E. M. Tomazou, N. P. Thorne, L. Backdahl, M. Herberth, K. L. Howe, D. K. Jackson, M. M. Miretti, J. C. Marioni, E. Birney, T. J. Hubbard, R. Durbin, S. Tavare and S. Beck, *Nature biotechnology*, 2008, **26**, 779-785.
67. G. Ficz, M. R. Branco, S. Seisenberger, F. Santos, F. Krueger, T. A. Hore, C. J. Marques, S. Andrews and W. Reik, *Nature*, 2011, **473**, 398-402.
68. E. A. Raiber, D. Beraldi, G. Ficz, H. E. Burgess, M. R. Branco, P. Murat, D. Oxley, M. J. Booth, W. Reik and S. Balasubramanian, *Genome biology*, 2012, **13**, R69.
69. L. Shen, H. Wu, D. Diep, S. Yamaguchi, A. C. D'Alessio, H. L. Fung, K. Zhang and Y. Zhang, *Cell*, 2013, **153**, 692-706.
70. D. Serre, B. H. Lee and A. H. Ting, *Nucleic acids research*, 2010, **38**, 391-399.
71. A. B. Brinkman, F. Simmer, K. Ma, A. Kaan, J. Zhu and H. G. Stunnenberg, *Methods*, 2010, **52**, 232-236.
72. S. Kangaspeska, B. Stride, R. Metivier, M. Polycarpou-Schwarz, D. Ibberson, R. P. Carmouche, V. Benes, F. Gannon and G. Reid, *Nature*, 2008, **452**, 112-115.
73. M. Mellen, P. Ayata, S. Dewell, S. Kriaucionis and N. Heintz, *Cell*, 2012, **151**, 1417-1430.
74. S. N. Kamdar, L. T. Ho, K. J. Kron, R. Isserlin, T. van der Kwast, A. R. Zlotta, N. E. Fleshner, G. Bader and B. Bapat, *Clinical epigenetics*, 2016, **8**, 32.
75. C. M. Greco, P. Kunderfranco, M. Rubino, V. Larcher, P. Carullo, A. Anselmo, K. Kurz, T. Carell, A. Angius, M. V. Latronico, R. Papait and G. Condorelli, *Nature communications*, 2016, **7**, 12418.
76. D. S. Johnson, A. Mortazavi, R. M. Myers and B. Wold, *Science*, 2007, **316**, 1497-1502.
77. R. Jothi, S. Cuddapah, A. Barski, K. Cui and K. Zhao, *Nucleic acids research*, 2008, **36**, 5221-5231.
78. G. Robertson, M. Hirst, M. Bainbridge, M. Bilenky, Y. Zhao, T. Zeng, G. Euskirchen, B. Bernier, R. Varhol, A. Delaney, N. Thiessen, O. L. Griffith, A. He, M. Marra, M. Snyder and S. Jones, *Nature methods*, 2007, **4**, 651-657.
79. H. S. Rhee and B. F. Pugh, *Cell*, 2011, **147**, 1408-1419.
80. D. A. Orlando, M. W. Chen, V. E. Brown, S. Solanki, Y. J. Choi, E. R. Olson, C. C. Fritz, J. E. Bradner and M. G. Guenther, *Cell reports*, 2014, **9**, 1163-1170.
81. I. Cheung, H. P. Shulha, Y. Jiang, A. Matevossian, J. Wang, Z. Weng and S. Akbarian, *Proceedings of the National Academy of Sciences of the United States of America*, 2010, **107**, 8824-8829.
82. Y. Kitagawa, N. Ohkura, Y. Kidani, A. Vandenbon, K. Hirota, R. Kawakami, K. Yasuda, D. Motooka, S. Nakamura, M. Kondo, I. Taniuchi, T. Kohwi-Shigematsu and S. Sakaguchi, *Nature immunology*, 2016, DOI: 10.1038/ni.3646.
83. R. Mathur, B. H. Alver, A. K. San Roman, B. G. Wilson, X. Wang, A. T. Agoston, P. J. Park, R. A. Shivdasani and C. W. Roberts, *Nature genetics*, 2016, DOI: 10.1038/ng.3744.
84. R. Hansel-Hertsch, D. Beraldi, S. V. Lensing, G. Marsico, K. Zyner, A. Parry, M. Di Antonio, J. Pike, H. Kimura, M. Narita, D. Tannahill and S. Balasubramanian, *Nature genetics*, 2016, **48**, 1267-1272.

85. S. Beyaz, J. H. Kim, L. Pinello, M. E. Xifaras, Y. Hu, J. Huang, M. A. Kerenyi, P. P. Das, R. A. Barnitz, A. Herault, R. Dogum, W. N. Haining, O. H. Yilmaz, E. Passegue, G. C. Yuan, S. H. Orkin and F. Winau, *Nature immunology*, 2016, DOI: 10.1038/ni.3644.
86. M. de Dieuleveult, K. Yen, I. Hmitou, A. Depaux, F. Boussouar, D. Bou Dargham, S. Jounier, H. Humbertclaude, F. Ribierre, C. Baulard, N. P. Farrell, B. Park, C. Keime, L. Carriere, S. Berlivet, M. Gut, I. Gut, M. Werner, J. F. Deleuze, R. Olaso, J. C. Aude, S. Chantalat, B. F. Pugh and M. Gerard, *Nature*, 2016, **530**, 113-116.
87. T. R. Butt, D. B. Jump and M. E. Smulson, *Proceedings of the National Academy of Sciences of the United States of America*, 1979, **76**, 1628-1632.
88. J. Allan, R. M. Fraser, T. Owen-Hughes and D. Keszenman-Pereyra, *Journal of molecular biology*, 2012, **417**, 152-164.
89. J. G. Henikoff, J. A. Belsky, K. Krassovsky, D. M. MacAlpine and S. Henikoff, *Proceedings of the National Academy of Sciences of the United States of America*, 2011, **108**, 18318-18323.
90. L. Song and G. E. Crawford, *Cold Spring Harbor protocols*, 2010, **2010**, pdb prot5384.
91. A. P. Boyle, S. Davis, H. P. Shulha, P. Meltzer, E. H. Margulies, Z. Weng, T. S. Furey and G. E. Crawford, *Cell*, 2008, **132**, 311-322.
92. H. H. He, C. A. Meyer, S. S. Hu, M. W. Chen, C. Zang, Y. Liu, P. K. Rao, T. Fei, H. Xu, H. Long, X. S. Liu and M. Brown, *Nature methods*, 2014, **11**, 73-78.
93. J. D. Buenrostro, P. G. Giresi, L. C. Zaba, H. Y. Chang and W. J. Greenleaf, *Nature methods*, 2013, **10**, 1213-1218.
94. J. D. Buenrostro, B. Wu, H. Y. Chang and W. J. Greenleaf, *Current protocols in molecular biology*, 2015, **109**, 21 29 21-29.
95. P. G. Giresi, J. Kim, R. M. McDaniell, V. R. Iyer and J. D. Lieb, *Genome research*, 2007, **17**, 877-885.
96. P. G. Giresi and J. D. Lieb, *Methods*, 2009, **48**, 233-239.
97. J. Mieczkowski, A. Cook, S. K. Bowman, B. Mueller, B. H. Alver, S. Kundu, A. M. Deaton, J. A. Urban, E. Larschan, P. J. Park, R. E. Kingston and M. Y. Tolstorukov, *Nature communications*, 2016, **7**, 11485.
98. V. I. Risca, S. K. Denny, A. F. Straight and W. J. Greenleaf, *Nature*, 2016, DOI: 10.1038/nature20781.
99. M. R. Corces, J. D. Buenrostro, B. Wu, P. G. Greenside, S. M. Chan, J. L. Koenig, M. P. Snyder, J. K. Pritchard, A. Kundaje, W. J. Greenleaf, R. Majeti and H. Y. Chang, *Nature genetics*, 2016, **48**, 1193-1203.
100. S. Wahl, A. Drong, B. Lehne, M. Loh, W. R. Scott, S. Kunze, P. C. Tsai, J. S. Ried, W. Zhang, Y. Yang, S. Tan, G. Fiorito, L. Franke, S. Guarrera, S. Kasela, J. Kriebel, R. C. Richmond, M. Adamo, U. Afzal, M. Ala-Korpela, B. Albetti, O. Ammerpohl, J. F. Apperley, M. Beekman, P. A. Bertazzi, S. L. Black, C. Blancher, M. J. Bonder, M. Brosch, M. Carstensen-Kirberg, A. J. de Craen, S. de Lusignan, A. Dehghan, M. Elkalaawy, K. Fischer, O. H. Franco, T. R. Gaunt, J. Hampe, M. Hashemi, A. Isaacs, A. Jenkinson, S. Jha, N. Kato, V. Krogh, M. Laffan, C. Meisinger, T. Meitinger, Z. Y. Mok, V. Motta, H. K. Ng, Z. Nikolakopoulou, G. Nteliopoulos, S. Panico, N. Pervjakova, H. Prokisch, W. Rathmann, M. Roden, F. Rota, M. A. Rozario, J. K. Sandling, C. Schafmayer, K. Schramm, R. Siebert, P. E. Slagboom, P. Soininen, L. Stolk, K. Strauch, E. S. Tai, L. Tarantini, B. Thorand, E. F. Tigchelaar, R. Tumino, A. G. Uitterlinden, C. van Duijn, J. B. van Meurs, P. Vineis, A. R. Wickremasinghe, C. Wijmenga, T. P. Yang, W. Yuan, A. Zhernakova, R. L. Batterham,

G. D. Smith, P. Deloukas, B. T. Heijmans, C. Herder, A. Hofman, C. M. Lindgren, L. Milani, P. van der Harst, A. Peters, T. Illig, C. L. Relton, M. Waldenberger, M. R. Jarvelin, V. Bollati, R. Soong, T. D. Spector, J. Scott, M. I. McCarthy, P. Elliott, J. T. Bell, G. Matullo, C. Gieger, J. S. Kooner, H. Grallert and J. C. Chambers, *Nature*, 2017, **541**, 81-86.

101. M. Mehrmohamadi, L. K. Mentch, A. G. Clark and J. W. Locasale, *Nature communications*, 2016, **7**, 13666.
102. M. Kulis, A. Merkel, S. Heath, A. C. Queiros, R. P. Schuyler, G. Castellano, R. Beekman, E. Raineri, A. Esteve, G. Clot, N. Verdaguer-Dot, M. Duran-Ferrer, N. Russinol, R. Vilarrasa-Blasi, S. Ecker, V. Pancaldi, D. Rico, L. Agueda, J. Blanc, D. Richardson, L. Clarke, A. Datta, M. Pascual, X. Agirre, F. Prosper, D. Alignani, B. Paiva, G. Caron, T. Fest, M. O. Muench, M. E. Fomin, S. T. Lee, J. L. Wiemels, A. Valencia, M. Gut, P. Flicek, H. G. Stunnenberg, R. Siebert, R. Kuppers, I. G. Gut, E. Campo and J. I. Martin-Subero, *Nature genetics*, 2015, **47**, 746-756.
103. A. M. Richter, S. Kiehl, N. Köger, J. Breuer, T. Stiewe and R. H. Dammann, *Clinical epigenetics*, 2017, **9**.
104. S. Tashiro, T. Handa, A. Matsuda, T. Ban, T. Takigawa, K. Miyasato, K. Ishii, K. Kugou, K. Ohta, Y. Hiraoka, H. Masukata and J. Kanoh, *Nature communications*, 2016, **7**, 10393.
105. A. Luoni, R. Massart, V. Nieratschker, Z. Nemoda, G. Blasi, M. Gilles, S. H. Witt, M. J. Suderman, S. J. Suomi, A. Porcelli, G. Rizzo, L. Fazio, S. Torretta, A. Rampino, A. Berry, P. Gass, F. Cirulli, M. Rietschel, A. Bertolino, M. Deuschle, M. Szyf and M. A. Riva, *Translational psychiatry*, 2016, **6**, e943.
106. G. E. Crawford, S. Davis, P. C. Scacheri, G. Renaud, M. J. Halawi, M. R. Erdos, R. Green, P. S. Meltzer, T. G. Wolfsberg and F. S. Collins, *Nature methods*, 2006, **3**, 503-509.
107. G. Zhang and S. Pradhan, *IUBMB life*, 2014, **66**, 240-256.
108. D. S. Paul, A. E. Teschendorff, M. A. Dang, R. Lowe, M. I. Hawa, S. Ecker, H. Beyan, S. Cunningham, A. R. Fouts, A. Ramelius, F. Burden, S. Farrow, S. Rowlston, K. Rehnstrom, M. Frontini, K. Downes, S. Busche, W. A. Cheung, B. Ge, M. M. Simon, D. Bujold, T. Kwan, G. Bourque, A. Datta, E. Lowy, L. Clarke, P. Flicek, E. Libertini, S. Heath, M. Gut, I. G. Gut, W. H. Ouwehand, T. Pastinen, N. Soranzo, S. E. Hofer, B. Karges, T. Meissner, B. O. Boehm, C. Cilio, H. Elding Larsson, A. Lernmark, A. K. Steck, V. K. Rakyan, S. Beck and R. D. Leslie, *Nature communications*, 2016, **7**, 13555.
109. R. L. Hood, L. C. Schenkel, S. M. Nikkel, P. J. Ainsworth, G. Pare, K. M. Boycott, D. E. Bulman and B. Sadikovic, *Scientific reports*, 2016, **6**, 38803.
110. H. Yan, S. Tian, S. L. Slager and Z. Sun, *Epigenomics*, 2016, **8**, 1239-1258.
111. U. M. Litzenburger, J. D. Buenrostro, B. Wu, Y. Shen, N. C. Sheffield, A. Kathiria, W. J. Greenleaf and H. Y. Chang, *Genome biology*, 2017, **18**, 15.
112. H. Easwaran, H. C. Tsai and S. B. Baylin, *Molecular cell*, 2014, **54**, 716-727.
113. E. C. Small, L. Xi, J. P. Wang, J. Widom and J. D. Licht, *Proceedings of the National Academy of Sciences of the United States of America*, 2014, **111**, E2462-2471.
114. J. D. Buenrostro, B. Wu, U. M. Litzenburger, D. Ruff, M. L. Gonzales, M. P. Snyder, H. Y. Chang and W. J. Greenleaf, *Nature*, 2015, **523**, 486-490.
115. J. Zhong, Z. Ye, S. W. Lenz, C. R. Clark, A. Bharucha, G. Farrugia, K. D. Robertson, Z. Zhang, T. Ordog and J. H. Lee, *BMC Genomics*, 2017, **18**, 985.
116. J. A. Dahl and P. Collas, *Nucleic acids research*, 2008, **36**, e15.

117. R. Blecher-Gonen, Z. Barnett-Itzhaki, D. Jaitin, D. Amann-Zalcenstein, D. Lara-Astiaso and I. Amit, *Nature protocols*, 2013, **8**, 539-554.
118. J. A. Dahl and P. Collas, *Stem Cells*, 2007, **25**, 1037-1046.
119. D. Lara-Astiaso, A. Weiner, E. Lorenzo-Vivas, I. Zaretsky, D. A. Jaitin, E. David, H. Keren-Shaul, A. Mildner, D. Winter, S. Jung, N. Friedman and I. Amit, *Science*, 2014, **345**, 943-949.
120. C. A. Aguilar and H. G. Craighead, *Nature nanotechnology*, 2013, **8**, 709-718.
121. J. F. Edd, D. Di Carlo, K. J. Humphry, S. Koster, D. Irimia, D. A. Weitz and M. Toner, *Lab on a chip*, 2008, **8**, 1262-1264.
122. Y. Zhang, Y. Zhu, B. Yao and Q. Fang, *Lab on a chip*, 2011, **11**, 1545-1549.
123. A. K. White, M. VanInsberghe, O. I. Petriv, M. Hamidi, D. Sikorski, M. A. Marra, J. Piret, S. Aparicio and C. L. Hansen, *Proceedings of the National Academy of Sciences of the United States of America*, 2011, **108**, 13999-14004.
124. S. C. Kim, I. C. Clark, P. Shahi and A. R. Abate, *Analytical chemistry*, 2017, DOI: 10.1021/acs.analchem.7b04050.
125. A. M. Streets, X. Zhang, C. Cao, Y. Pang, X. Wu, L. Xiong, L. Yang, Y. Fu, L. Zhao, F. Tang and Y. Huang, *Proceedings of the National Academy of Sciences of the United States of America*, 2014, **111**, 7048-7053.
126. A. M. Klein, L. Mazutis, I. Akartuna, N. Tallapragada, A. Veres, V. Li, L. Peshkin, D. A. Weitz and M. W. Kirschner, *Cell*, 2015, **161**, 1187-1201.
127. Y. Fu, C. Li, S. Lu, W. Zhou, F. Tang, X. S. Xie and Y. Huang, *Proceedings of the National Academy of Sciences of the United States of America*, 2015, **112**, 11923-11928.
128. A. R. Wu, J. B. Hiatt, R. Lu, J. L. Attema, N. A. Lobo, I. L. Weissman, M. F. Clarke and S. R. Quake, *Lab on a chip*, 2009, **9**, 1365-1370.
129. A. R. Wu and S. R. Quake, *Cold Spring Harbor protocols*, 2016, **2016**, pdb prot084996.
130. A. R. Wu, T. L. Kawahara, N. A. Rapicavoli, J. van Riggelen, E. H. Shroff, L. Xu, D. W. Felsher, H. Y. Chang and S. R. Quake, *Lab on a chip*, 2012, **12**, 2190-2198.
131. Z. Cao, C. Chen, B. He, K. Tan and C. Lu, *Nature methods*, 2015, **12**, 959-962.
132. Z. Cao and C. Lu, *Analytical chemistry*, 2016, **88**, 1965-1972.
133. T. W. Murphy, Y. P. Hsieh, S. Ma, Y. Zhu and C. Lu, *Analytical chemistry*, 2018, **90**, 7666-7674.
134. A. Rotem, O. Ram, N. Shoresh, R. A. Sperling, A. Goren, D. A. Weitz and B. E. Bernstein, *Nature biotechnology*, 2015, **33**, 1165-1172.
135. L. R. Cardon and T. Harris, *Human Molecular Genetics*, 2016, **25**, R166-R172.
136. S. Y. Teh, R. Lin, L. H. Hung and A. P. Lee, *Lab on a chip*, 2008, **8**, 198-220.
137. R. Seemann, M. Brinkmann, T. Pfohl and S. Herminghaus, *Rep Prog Phys*, 2012, **75**, 016601.
138. O. J. Dressler, R. M. Maceiczyk, S. I. Chang and A. J. deMello, *Journal of biomolecular screening*, 2014, **19**, 483-496.
139. T. S. Kaminski, O. Scheler and P. Garstecki, *Lab on a chip*, 2016, **16**, 2168-2187.
140. D.-K. Kang, M. Monsur Ali, K. Zhang, E. J. Pone and W. Zhao, *TrAC Trends in Analytical Chemistry*, 2014, **58**, 145-153.
141. C. N. Baroud, F. Gallaire and R. Dangla, *Lab on a chip*, 2010, **10**, 2032-2045.
142. G. F. Christopher and S. L. Anna, *Journal of Physics D: Applied Physics*, 2007, **40**, R319-R336.
143. M. Leman, F. Abouakil, A. D. Griffiths and P. Tabeling, *Lab on a chip*, 2015, **15**, 753-765.

144. D. T. Chiu, R. M. Lorenz and G. D. Jeffries, *Analytical chemistry*, 2009, **81**, 5111-5118.
145. M. Chabert and J. L. Viovy, *Proceedings of the National Academy of Sciences of the United States of America*, 2008, **105**, 3191-3196.
146. H. H. Gorris and D. R. Walt, *Angew Chem Int Ed Engl*, 2010, **49**, 3880-3895.
147. E. W. Kemna, R. M. Schoeman, F. Wolbers, I. Vermes, D. A. Weitz and A. van den Berg, *Lab on a chip*, 2012, **12**, 2881-2887.
148. D. J. Eastburn, A. Sciambi and A. R. Abate, *Analytical chemistry*, 2013, **85**, 8016-8021.
149. T. M. Tran, F. Lan, C. S. Thompson and A. R. Abate, *Journal of Physics D: Applied Physics*, 2013, **46**, 114004.
150. P. Garstecki, M. J. Fuerstman, H. A. Stone and G. M. Whitesides, *Lab on a chip*, 2006, **6**, 437-446.
151. H. N. Joensson and H. Andersson Svahn, *Angew Chem Int Ed Engl*, 2012, **51**, 12176-12192.
152. L. Mazutis, J. Gilbert, W. L. Ung, D. A. Weitz, A. D. Griffiths and J. A. Heyman, *Nature protocols*, 2013, **8**, 870-891.
153. T. P. Lagus and J. F. Edd, *Journal of Physics D: Applied Physics*, 2013, **46**, 114005.
154. A. Rakszewska, J. Tel, V. Chokkalingam and W. T. S. Huck, *NPG Asia Materials*, 2014, **6**, e133.
155. A. R. Abate, C. H. Chen, J. J. Agresti and D. A. Weitz, *Lab on a chip*, 2009, **9**, 2628-2631.
156. M. Y. He, J. S. Edgar, G. D. M. Jeffries, R. M. Lorenz, J. P. Shelby and D. T. Chiu, *Analytical chemistry*, 2005, **77**, 1539-1544.
157. L. Mazutis and A. D. Griffiths, *Lab on a chip*, 2012, **12**, 1800-1806.
158. Y. Ding, X. C. i Solvas and A. deMello, *The Analyst*, 2015, **140**, 414-421.
159. A. Sciambi and A. R. Abate, *Lab on a chip*, 2015, **15**, 47-51.
160. V. Chokkalingam, Y. Ma, J. Thiele, W. Schalk, J. Tel and W. T. Huck, *Lab on a chip*, 2014, **14**, 2398-2402.
161. E. Brouzes, A. Carniol, T. Bakowski and H. H. Strey, *RSC advances*, 2014, **4**, 38542-38550.
162. Y.-C. Tan, Y. L. Ho and A. P. Lee, *Microfluidics and Nanofluidics*, 2006, **3**, 495-499.
163. D. R. Link, S. L. Anna, D. A. Weitz and H. A. Stone, *Physical Review Letters*, 2004, **92**.
164. B. O'Donovan, T. Tran, A. Sciambi and A. Abate, *Journal of visualized experiments : JoVE*, 2014, DOI: 10.3791/50913.
165. L. M. Fidalgo, C. Abell and W. T. Huck, *Lab on a chip*, 2007, **7**, 984-986.
166. L. Li, J. Q. Boedicker and R. F. Ismagilov, *Analytical chemistry*, 2007, **79**, 2756-2761.
167. A. R. Abate, T. Hung, P. Mary, J. J. Agresti and D. A. Weitz, *Proceedings of the National Academy of Sciences of the United States of America*, 2010, **107**, 19163-19166.
168. S. R. Doonan and R. C. Bailey, *Analytical chemistry*, 2017, **89**, 4091-4099.
169. P. Tabeling, *Lab on a chip*, 2009, **9**, 2428-2436.
170. S. Ma, J. M. Sherwood, W. T. Huck and S. Balabani, *Lab on a chip*, 2014, **14**, 3611-3620.
171. J. D. Tice, H. Song, A. D. Lyon and R. F. Ismagilov, *Langmuir*, 2003, **19**, 9127-9133.

Chapter 2 Design, Fabrication, and Operation of Droplet Microfluidic Devices

Introduction

One of the most common methods to generate prototypes of microfluidic devices in academic laboratories is soft lithography with poly(dimethylsiloxane) (PDMS). Ever since Whitesides' group reported the successful fabrication of microfluidic channels with this technique in the 90s[1], this rapid prototyping protocol with PDMS for microfluidic devices has obtained wide applications due to its favorable features including optical transparency, high biocompatibility, gas permeability, low cost, and ease of use.[2, 3] This chapter describes the adapted protocol of soft lithography with PDMS for droplet microfluidic device fabrication for most of the work in this thesis in detail. The process of device fabrication in following chapters is the same as what is described here unless otherwise noted.

Materials and Methods

Materials

Isopropanol

Negative photoresist SU8 2025 and 2050 (MicroChem Corp, Westborough, MA)

Propylene glycol monomethyl ether acetate (PGMEA) (Sigma-Aldrich, St. Louis, MO)

(Tridecafluoro-1,1,2,2-tetrahydrooctyl) trichlorosilane (Gelest, Inc., Morrisville, PA)

PDMS base and curing agent (RTV 615 kit, Momentive Performance Materials Inc., Water Ford, NY)

Fluorinert FC40 (Sigma-Aldrich, St. Louis, MO)

Aquapel (Pittsburgh Glass Works, LLC., Pittsburgh, PA)

Novec 7500 engineered fluid (3M, St. Paul, MN)

Surfactant (Perfluoropolyether polyethylene glycol block co-polymer, RAN Biotechnologies, Inc. Beverly, MA)

Sodium chloride solution (3 mole/L, aqueous solution)

Silicon wafers (University Wafer, Boston MA)

UV lamp (Optical Associates, Incorporated)

Hot plates (Cimarec, Thermo Scientific)

Oven (HeraTherm Oven, Thermo Scientific)

Spin coater (PWM 32, Headway Research, Inc)

Long pass UV filter (PL360-LP, Omega Optical, Inc., Brattleboro, VT)

Profilometer (KLA-tencor Alpha-Step IQ or Dektak XT Surface Profilometers)

Forceps or tweezers

Transfer pipettes

Glass petri dishes

Plastic petri dishes

Blunt-tip biopsy needles

Scotch Magic Invisible Tapes (3M, St. Paul, MN)

Syringe filters w/ 0.2 μm PTFE membrane (VWR International, Inc)

Microscope glass slide (Fisher Scientific, Inc.)

Power supply (DC regulated power supply, TENMA T2-6628)

In-house built DC-AC converter (12V DC input, 0-360V AC at 36 kHz as output)

Syringe pumps (Pump 11 Pico Plus Elite, Harvard Apparatus)

High speed camera (Phantom Miro Ex2 or Phantom VEO 640L)

Stereoscope or fluorescence microscope (Leica M80 or DMi8)

Design and Fabrication of Masters

Microfluidic patterns were designed with a computer-aided design (CAD) software program AutoCAD (AutoDesk, Inc.) and the cad files were sent to CAD/Art Services, Inc. (Bandon, OR) for photomask printing with the polarity of patterns being transparent and other areas covered with black ink (See **Figure A-1** in **Appendix A**). Once photomasks were received, standard contact photolithography was conducted to generate masters.[4] Depending on the thickness of the designed patterns and the effective power of the UV lamp, fabrication parameters including spin rate, spin duration, baking temperatures, and UV exposure duration were determined accordingly following the guidance of the photoresist vendor datasheet and testing trials. To obtain clean masters, it was preferred to execute the protocol in a cleanroom. A detailed fabrication process was described as below to obtain a master for the nucleosome preparation device used in Chapter 3 that contained patterns of different thickness (the pattern for the droplet formation and manipulation was referred to as the first layer, and the pattern for droplet delay was referred to as the second layer). A general workflow is shown in **Figure 2-1**.

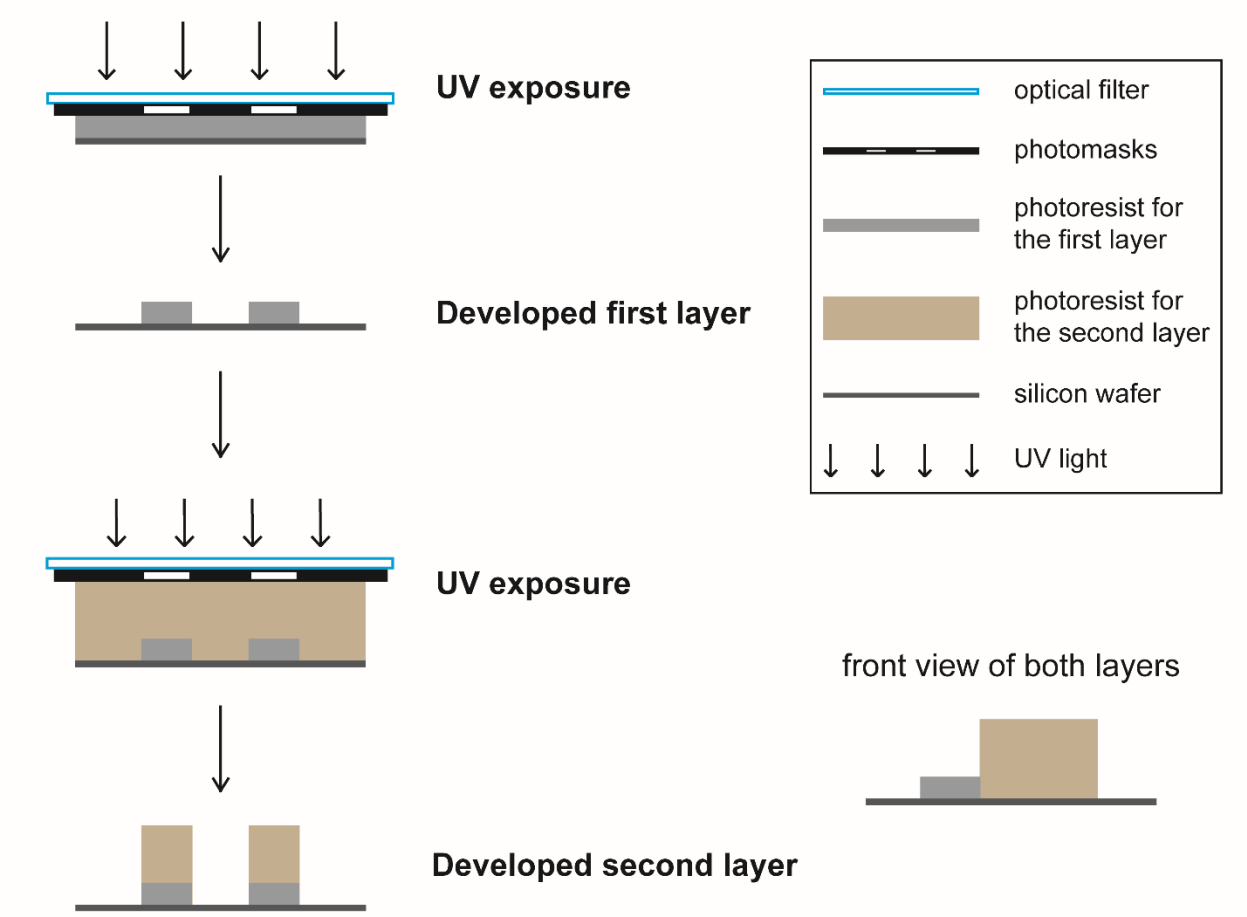

Figure 2-1 Fabrication workflow of microfluidic device masters using negative photoresist to construct two layers of patterns that are of different thicknesses.

A silicon wafer was dried at 120 °C for 5 minutes. A quarter-sized SU8 2025 was aliquoted onto a 3-inch silicon wafer with a transfer pipette, and the wafer was centered onto the holder of a spin coater. A spin rate of 500 rpm for the duration of 10 seconds followed by a spin rate of 2000 rpm for the duration of 30 seconds was operated to achieve a 40-µm layer. After soft baking at 65 °C for 3 minutes and 95 °C for 6 minutes, respectively, the silicon wafer was set aside to cool down for 5 minutes. The first-layer photomask was aligned to the wafer, and a filter to cut off unnecessary wavelengths of light (<350 nm) was placed on top of the mask. Two weights made of iron were put at the edge of the filter to help tighten the contact of the mask to the photoresist. UV

exposure was completed with a UV lamp (Optical Associates, Incorporated) for 24 seconds (**Figure 2-1, the first "UV exposure"**). After the removal of the filter and photomask, the silicon wafer was set aside to rest for 5 minutes and transferred with forceps to complete post-exposure baking at 65 °C for 1 minutes and 95 °C for 6 minutes, respectively. After set aside and cool for 5 minutes, the silicon wafer was immersed into propylene glycol monomethyl ether acetate (PGMEA) in a glass petri dish for 5 to 10 minutes to develop the uncured photoresist. During the development step, gentle agitation of the petri dish or the silicon wafer was favored to accelerate the dissolution of uncured SU8 2025. The wafer was rinsed thoroughly with fresh PGMEA and with isopropanol to ensure that no white cloudy material was formed, indicating that all uncured photoresist was removed. Then the wafer was blow dry with nitrogen and set aside to rest for 5 minutes. The first layer was completed (**Figure 2-1, "Developed first layer"**).

A quarter-sized SU8-2050 was aliquoted onto the wafer and spin coated at 500 rpm for 10 seconds followed by 1150 rpm for 30 seconds to achieve a 160-μm layer. Then the wafer was soft baked at 65 °C for 7 min and 95 °C for 35 minutes, respectively. After set aside to cool down for 5 minutes, the second-layer photomask was aligned (via visual estimation or a professional mask aligner with the assistance of overlapped patterns and/or align markers) to the existing pattern to ensure a correct connection of channels. Then the photomask was tightly contacted to the photoresist with the addition of the filter and iron weights. UV exposure was operated for 36 seconds (**Figure 2-1, the second "UV exposure"**) and the wafer was set aside to cool down for 5 minutes followed by post-exposure baking at 65 °C for 5 minutes and 95 °C for 15 minutes, respectively. After another 5 minutes of cooling down, the photoresist was developed in PGMEA, blow dry with nitrogen, and set aside for cool down as described before. The second layer was completed (**Figure 2-1**, **"Developed second layer"**).

The fabricated master was placed onto a room-temperature hot plate, which was then set to 200 °C. The master was kept on the hot plate for 10 min after it reached the targeted temperature. After that, the hot place was turned off and the wafer remained on the hot plate until it naturally cooled down to room temperature to avoid sudden thermal change induced contraction of SU8 and subsequent detachment of patterns. This step of hard baking was supposed to heal any potential cracks on the solidified SU8 patterns as well. The master was then placed in a plastic petri dish and put into a vacuum chamber to be treated with vacuum-assisted evaporated 100-µL (tridecafluoro-1,1,2,2-tetrahydrooctyl) trichlorosilane overnight.

To check the fabricated patterns' real thicknesses, a profilometer can be utilized to measure the height of interested planes. The profilometer's stylus tip moves vertically to get in contact with a surface and scans over the surface with a preset contact force to record the vertical displacement variation cross a certain lateral distance. A profile like **Figure 2-2** was constructed based on the scanned data showing the thickness of the first layer to be approximately 46 µm and the difference between the first and second layers to be approximately 94 µm, indirectly measuring the thickness of the second layer to be approximately 140 µm. The fabricated patterns had similar thicknesses compared to the theoretical values, indicating a quality controlled, successful master fabrication.

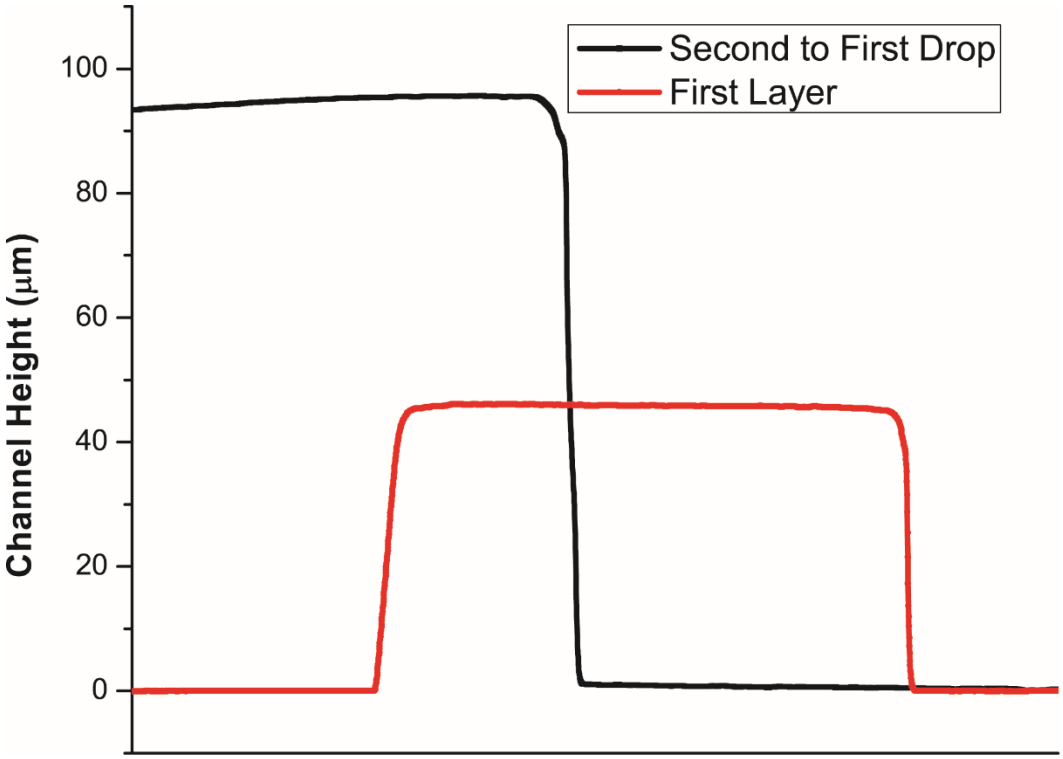

Figure 2-2 Profilometer measurement of a master with patterns of two different thicknesses.

Assembly of PDMS devices

The microfluidic devices were fabricated by soft lithography as described in previous literature[1, 5] with slight adjustment. A general workflow is shown in **Figure 2-3**. Briefly, PDMS base and curing agent were mixed at a mass ratio of 10:1 and degassed under vacuum before the mixture was poured over the fabricated device master in the last section (**Figure 2-3, "Pour PDMS to form a stamp"**). After being cured in an oven at 70 °C for at least 1 hour, PDMS stamps were cut and peeled off from the master. Holes of inlets and outlets were punched with blunt-tip biopsy needles and the stamps were sonicated clean in water for 5 min (**Figure 2-3, "Peel off stamp to punch holes for tubing"**). They were dried with nitrogen and cleaned with Scotch tapes. A degassed mixture of PDMS base and curing agent at a mass ratio of 10:1 was spread over a microscope slide on the spin coater at 500 rpm for 15 seconds followed by 1500 rpm for 30 seconds

and was subsequently half-cured in the oven at 70 °C for approximately 8 min (the duration should be empirically determined to ensure half cure of PDMS). Subsequently, the cleaned PDMS stamp was gently attached to the PDMS coated glass slide with mild pressing from a tweezer. The assembled device was then baked in the oven at 70 °C for at least 30 min before use (**Figure 2-3, "Assemble device onto a flat surface"**).

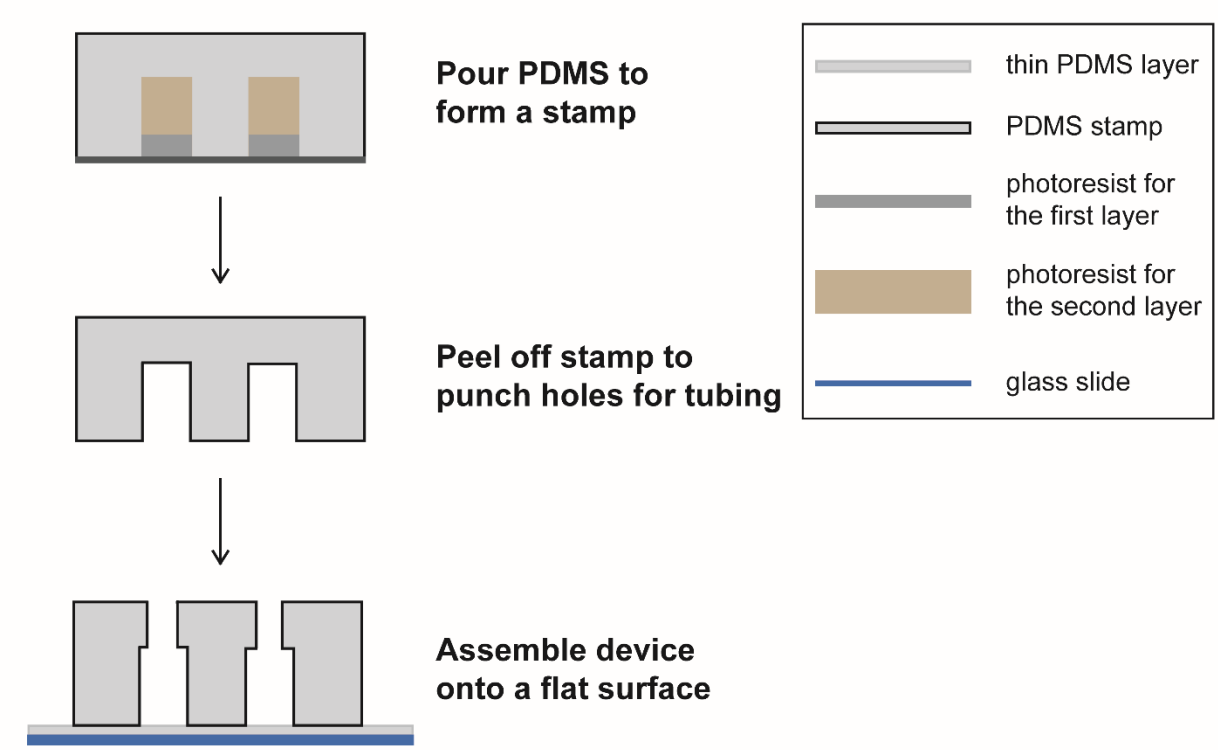

Figure 2-3 Workflow of microfluidic device fabrication by soft lithography using PDMS.

Employment of Assembled Devices

A PDMS device, when ready to use for experiments, was taken out of the oven and surfaced treated with Aquapel as described in previous literatures.[6] Briefly the device was filled and incubated with Aquapel for ~ 60 seconds. Aquapel was expelled out the device by air. Then the device was rinsed with filtered FC40. Electrolyte channels were filled with sodium chloride solution (3 M). Two clamps from electrodes of an in-house built DC-AC converter (12V DC input,

0-360V AC at 36 kHz as output) which was connected to a DC-regulated power supply were connected to the syringe needles supplying the electrolyte channels to form a complete circuit for the AC electric field. The device is placed on the holder of a stereoscope (Leica M80) or a fluorescence microscope (Leica DMi8).

The components of the reagents delivered into the device depend on the specific experiments that are to be operated. Generally, aqueous reagents for droplet formation are supplied by 24G PTFE tubing (Cole-Parmer, Vernon Hills, IL) attached to syringes driven by one syringe pump. Novec 7500 engineered fluid, the fluorinated oil phase for droplet formation, is mixed with 2% (w/w) surfactant and delivered to the device by a second syringe pump. Other reagents for additional manipulations of droplets are delivered to the device by additional pumps. Flow rates are optimized to ensure that droplets with appropriate size were generated and subsequent manipulations are carried out successfully. Droplets are monitored by a high-speed camera mounted onto the stereoscope or the fluorescence microscope. Products are collected to an Eppendorf tube for subsequent processing.

Operations of Droplet Microfluidics

Examples of basic droplet microfluidic operations that are utilized in this thesis include droplet formation, droplet delay, and reagent addition. Monodispersed droplets can be continuously generated at a frequency ranging from 10^0 Hz to > 10^3 Hz using common formation geometries such as a T-junction[7] and a flow focusing[8, 9] design. Demonstration of droplet formation were shown in **Figure 2-4a and b** with in-house built T-junction and flow focusing devices. The integrity of droplets is maintained by the protection from the surfactant that is doped in the continuous phase (oil) and adsorbed at the interface between droplets and the oil.[10-12] This

enables droplets to travel for a certain duration without coalescence, and thus allows the completion of the physical, chemical and/or biological process inside (**Figure 2-4c**). Under the effect of an electric field, the dynamic instability of the surfactant layer at the interface between droplets and the continuous phase allows the addition of another stream of aqueous phase into the otherwise protected droplets (**Figure 2-4d**).[13-15]

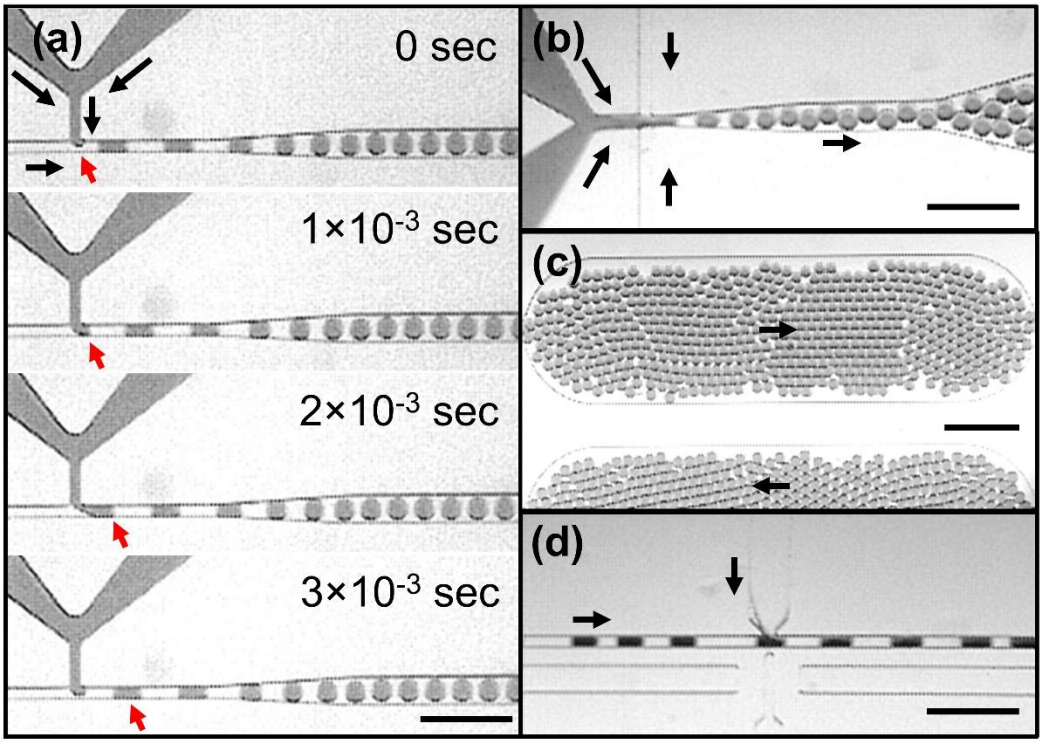

Figure 2-4 Examples of droplet microfluidic operations at 8 µL/min of the oil and 2 µL/min of each aqueous stream for droplet formation using in-house built devices. (a) Time sequence of droplet formation in a T-junction geometry. Notice that the formation of one droplet took less than 3×10-3 sec, indicating a formation frequency higher than 333 Hz. The black arrows indicated the liquid flow directions, and the red arrows indicated one specific droplet. (b) Droplet formation in a flow-focusing geometry. (c) Droplets remained the same in two delay channels that were separated by multiple others. (d) Reagent addition via the geometry of picoinjector into every single droplet. The flow rate of the injected stream was 1.2 µL/min. Scale bar represents 300 µm (a, b, and d) and 500 µm (c), respectively.

Previous studies have been trying to elucidate the mechanism and governing rules of droplet formation, and they have revealed that the flow rates and the viscosities of the participating

liquids[8, 16-19], the dimensions of the channels[18, 20, 21], and the type and the concentration of the utilized surfactant[12, 22] all contribute to the final morphology and generating mechanism of the produced droplets. In the case of a T-junction used under typical microfluidic conditions (flow rates on the order of 0.01 to 1 μL/s and the capillary number smaller than 10^{-2}) and channel widths and heights on the order of 10 to 100 μm, a simple scaling rule (2.1) can be used to predict droplet sizes, where l is the length of the droplet (that blocks the forming channel), w is the width of the channel, Q_d and Q_c are flow rates of the dispersed and continuous phases respectively, and α is a constant that depends on the geometry of the T-junction.[18] As shown in **Figure 2-5**, increasing the flow rate of oil decreased the flow rate ratio between the dispersed and continuous phases when the dispersed phase were delivered to the device at a constant total flow rate, leading to a reduction in the length of the droplets generated by the in-house T-junction device with a channel width of 40 μm. The linear fitting yielded a high regression coefficient ($R^2 = 0.999$) at a fixed intercept of 40, evidenced that the in-house experimental setup of droplet generation with a T-junction has quantitatively matched with precedent literature results.

$$\frac{l}{w} = 1 + \frac{\alpha Q_d}{Q_c} \tag{2.1}$$

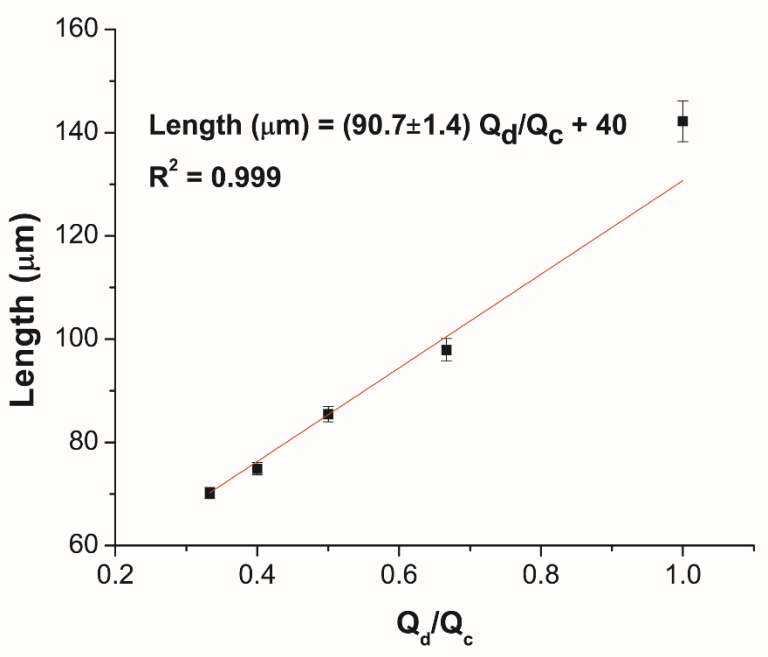

Figure 2-5 Adjusting droplet sizes by changing the flow rates of participating liquids. Q_d was kept constant at 4 µL/min (2 µL/min of each aqueous stream coming to the formation junction) while *Qc* was changed from 4 to 12 µL/min at a 2 µL/min step. n ≥ 18 droplets for all conditions.

Similarly, changing flow rates of the liquid also modulates the droplet formation rate when using the same design of T-junction.[23] Increasing Q_d or Q_c accelerates the production rate of droplets, which is intuitively obvious as more materials are delivered to the same channel per unit of time (**Figure 2-6**). Notice that the production rate was on the order of 10^2 to 10^3 Hz and the generated droplet size was on the order of 10^2 pL, demonstrating the potential for droplet microfluidics to be exploited in the fields where high throughput and small, individual encapsulation are desired including high throughput screening[24, 25] and single cell studies[26-28].

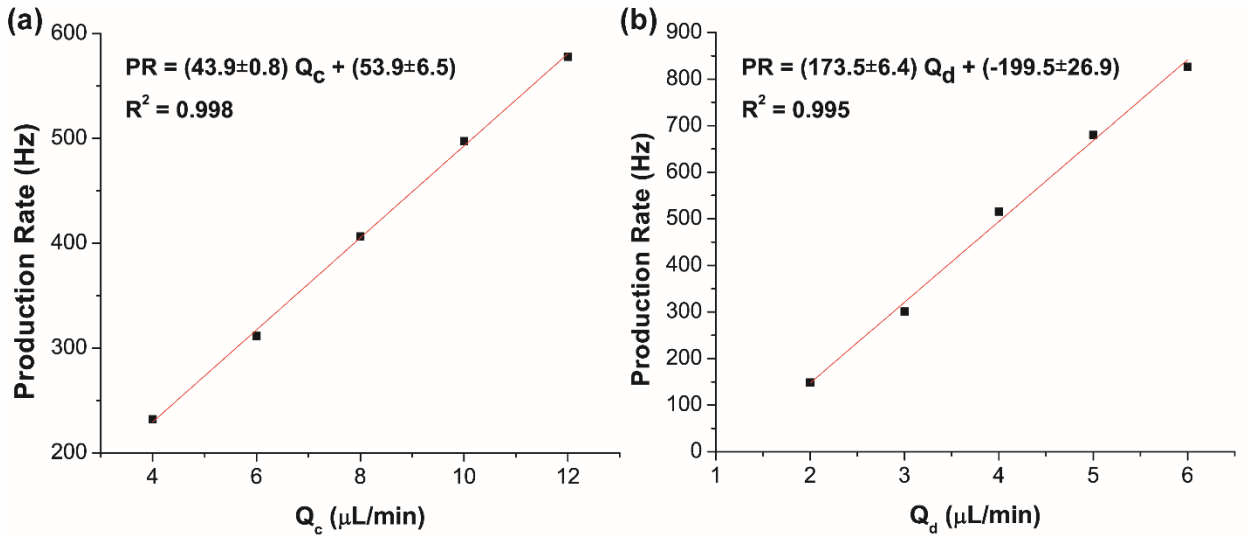

Figure 2-6 Production rate of droplets under different flow rates of the continuous and dispersed phases. (a) The Q_d of the aqueous streams was kept constant at 4 µL/min. (b) The flow rate ratio between the dispersed and continuous phases was maintained at 2. n ≥ 18 droplets for all conditions.

Not only the sizes and the production rate of droplets can be controlled, the amount of material injected into each droplet can also be tuned when interdroplet spacing is the same for all droplets coming across the injecting site as an array.[14] Increasing the flow rate of the to-be-injected aqueous stream while maintaining other conditions constant increased the amount of fluid added to droplets (**Figure 2-7**).

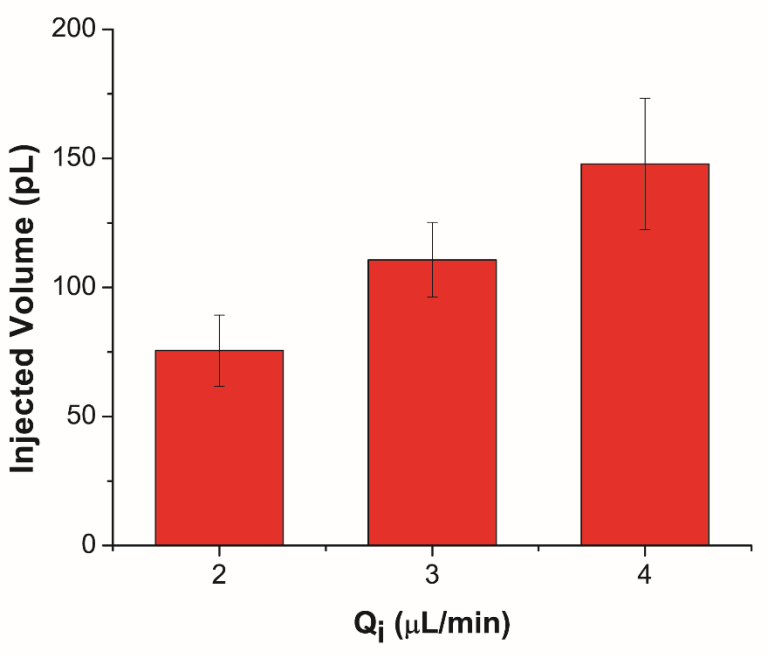

Figure 2-7 Injected volume of the to-be-injected aqueous stream under different flow rates (Q_i). The flow rates of the dispersed (Q_d) and continuous (Q_c) phases for droplet formation were maintained at 6 and 8 μL/min, respectively. n ⩾ 43 droplets for all conditions.

Conclusion

In this chapter, a complete workflow of implementing droplet microfluidic experiments has been introduced from the design of microfluidic device patterns to the common operations utilizing the devices. The presented data proved that in-house built devices fabricated with standard photolithography and soft lithography can generate expected results that have been established by previous literatures following the protocol explained here. Droplet formation, delay, and reagent injection were carried out successfully with common geometries including T-junctions, flow focusing, and picoinjectors. Manipulating droplets carefully can be significant to control parameters that can contribute to background signals when developing quantitative assays while a wider range of droplet sizes, injected volumes, and other parameters are acceptable for other applications.[29, 30] Therefore, the extent to adjust droplets depend on the applications, and the

protocol described in this chapter can serve as a general guidance to initiate droplet microfluidic device design and tests for subsequent assay development.

References

1. D. C. Duffy, J. C. McDonald, O. J. Schueller and G. M. Whitesides, *Analytical chemistry*, 1998, **70**, 4974-4984.
2. D. Qin, Y. Xia and G. M. Whitesides, *Nature protocols*, 2010, **5**, 491-502.
3. X. Hou, Y. S. Zhang, G. T.-d. Santiago, M. M. Alvarez, J. Ribas, S. J. Jonas, P. S. Weiss, A. M. Andrews, J. Aizenberg and A. Khademhosseini, *Nature Reviews Materials*, 2017, **2**, 17016.
4. R. Seemann, M. Brinkmann, T. Pfohl and S. Herminghaus, *Rep Prog Phys*, 2012, **75**, 016601.
5. J. C. McDonald, D. C. Duffy, J. R. Anderson, D. T. Chiu, H. Wu, O. J. A. Schueller and G. M. Whitesides, *Electrophoresis*, 2000, **21**, 27-40.
6. L. Mazutis, J. Gilbert, W. L. Ung, D. A. Weitz, A. D. Griffiths and J. A. Heyman, *Nature protocols*, 2013, **8**, 870-891.
7. T. Thorsen, R. W. Roberts, F. H. Arnold and S. R. Quake, *Physical review letters*, 2001, **86**, 4163-4166.
8. S. L. Anna, N. Bontoux and H. A. Stone, *Applied Physics Letters*, 2003, **82**, 364-366.
9. R. Dreyfus, P. Tabeling and H. Willaime, *Physical review letters*, 2003, **90**, 144505.
10. J. C. Baret, *Lab on a chip*, 2012, **12**, 422-433.
11. B. Riechers, F. Maes, E. Akoury, B. Semin, P. Gruner and J. C. Baret, *Proceedings of the National Academy of Sciences of the United States of America*, 2016, **113**, 11465-11470.
12. N. M. Kovalchuk, E. Roumpea, E. Nowak, M. Chinaud, P. Angeli and M. J. H. Simmons, *Chemical Engineering Science*, 2018, **176**, 139-152.
13. M. Chabert, K. D. Dorfman and J. L. Viovy, *Electrophoresis*, 2005, **26**, 3706-3715.
14. A. R. Abate, T. Hung, P. Mary, J. J. Agresti and D. A. Weitz, *Proceedings of the National Academy of Sciences of the United States of America*, 2010, **107**, 19163-19166.
15. B. O'Donovan, T. Tran, A. Sciambi and A. Abate, *Journal of visualized experiments : JoVE*, 2014, DOI: 10.3791/50913.
16. A. Gupta, S. M. S. Murshed and R. Kumar, *Applied Physics Letters*, 2009, **94**, 164107.
17. S. van der Graaf, T. Nisisako, C. G. Schroen, R. G. van der Sman and R. M. Boom, *Langmuir*, 2006, **22**, 4144-4152.
18. P. Garstecki, M. J. Fuerstman, H. A. Stone and G. M. Whitesides, *Lab on a chip*, 2006, **6**, 437-446.
19. J. D. Tice, A. D. Lyon and R. F. Ismagilov, *Analytica Chimica Acta*, 2004, **507**, 73-77.
20. W. Zeng, S. Li and H. Fu, *Sensors and Actuators A: Physical*, 2018, **272**, 11-17.
21. W. Lee, L. M. Walker and S. L. Anna, *Physics of Fluids*, 2009, **21**, 032103.
22. J. C. Baret, F. Kleinschmidt, A. El Harrak and A. D. Griffiths, *Langmuir*, 2009, **25**, 6088-6093.
23. H. Fu, W. Zeng and S. Li, *Journal of Micromechanics and Microengineering*, 2017, **27**, 125020.

24. M. M. Kiss, L. Ortoleva-Donnelly, N. R. Beer, J. Warner, C. G. Bailey, B. W. Colston, J. M. Rothberg, D. R. Link and J. H. Leamon, *Analytical chemistry*, 2008, **80**, 8975-8981.
25. M. Sesen, T. Alan and A. Neild, *Lab on a chip*, 2017, **17**, 2372-2394.
26. H. N. Joensson and H. Andersson Svahn, *Angew Chem Int Ed Engl*, 2012, **51**, 12176-12192.
27. D.-K. Kang, M. Monsur Ali, K. Zhang, E. J. Pone and W. Zhao, *TrAC Trends in Analytical Chemistry*, 2014, **58**, 145-153.
28. A. M. Klein and E. Macosko, *Lab on a chip*, 2017, DOI: 10.1039/c7lc90070h.
29. F. W. Liu, S. T. Ding, E. C. Lin, Y. W. Lu and J. S. R. Jang, *RSC advances*, 2017, **7**, 4646-4655.
30. J. D. Tice, H. Song, A. D. Lyon and R. F. Ismagilov, *Langmuir*, 2003, **19**, 9127-9133.

Chapter 3 A Droplet Microfluidic Platform for Efficient Enzymatic Chromatin Digestion Enables Robust Determination of Nucleosome Positioning

This chapter has been adapted with permission from The Royal Society of Chemistry from the article "A Droplet Microfluidic Platform for Efficient Enzymatic Chromatin Digestion Enables Robust Determination of Nucleosome Positioning". The original article (Y. Xu, J.-H. Lee, Z. Li, L. Wang, T. Ordog and R. C. Bailey, Lab on a chip, 2018, DOI: 10.1039/c8lc00599k) can be accessed online at http://pubs.rsc.org/en/content/articlehtml/2018/lc/c8lc00599k.

This work was funded by the National Cancer Institute via CA191186, the Mayo-Illinois Alliance for Technology Based Healthcare, the Mayo Clinic Center for Individualized Medicine, and the NIH-sponsored Midwest Cancer Nanotechnology Training Center (CA154015). I sincerely appreciate the discussion regarding assay development with Dr. Jeong-Heon Lee (Epigenomics Program, Center for individualized Medicine, Mayo Clinic, Rochester, Minnesota, USA). I also gratefully appreciate the assistance in bioinformatic analysis from Dr. Zhaoyu Li (Department of Cancer Biology, Mayo Clinic, Jacksonville, Florida, USA), Dr. Liguo Wang (Department of Health Sciences Research, Mayo Clinic, Rochester, Minnesota, USA), and Dr. Yongbing Zhao (Department of Cancer Biology, Mayo Clinic, Jacksonville, Florida, USA).

Abstract

The first step in chromatin-based epigenetic assays involves the fragmentation of chromatin to facilitate precise genomic localization of the associated DNA. Here, we report the

development of a droplet microfluidic device that can rapidly and efficiently digest chromatin into single nucleosomes starting from whole-cell input material offering simplified and automated processing compared to conventional manual preparation. We demonstrate the digestion of chromatin from 2,500-125,000 Jurkat cells using micrococcal nuclease for enzymatic processing. We show that the yield of mononucleosomal DNA can be optimized by controlling enzyme concentration and incubation time, with resulting mononucleosome yields exceeding 80%. Bioinformatic analysis of sequenced mononucleosomal DNA (MNase-seq) indicated a high degree of reproducibility and concordance (97-99%) compared with conventionally processed preparations. Our results demonstrate the feasibility of robust and automated nucleosome preparation using a droplet microfluidic platform for nucleosome positioning and downstream epigenomic assays.

Introduction

In eukaryotic organisms, meters of genomic DNA must be compacted to fit into a micron-sized cell nucleus.[1, 2] This is accomplished by the formation of nucleoprotein complexes—termed chromatin for their readily stainable nature—that permit high packing ratios. However, this arrangement must still allow for controlled access to the DNA for template-dependent functions such as gene transcription, DNA replication, and repair.[3-6] Dynamic and highly regulated compaction and decompaction of DNA reflects changes in the density of nucleosomes, the fundamental repeating unit of chromatin. With ~147 base pairs (bp) of DNA coiling around an octamer of histones H2A, H2B, H3 and H4, loosely organized nucleosomes form a beads-on-a-string array structure connected by linker DNA.[3-6] Nucleosomes usually occupy positions on DNA sequences that are correlated to specific functions.[2, 3, 7] They are commonly depleted from active

promoter, enhancer, or terminator regions, forming nucleosome-depleted regions (NDRs) or nucleosome-free regions (NFRs) that allow for easier access of transcriptional machineries to these genes[8, 9]; while in intragenic regions nucleosomes have a more random and continuous distribution that is proposed to assist in maintaining genome stability by occluding DNA-binding proteins.[4, 10] Dynamic nucleosome repositioning occurs in response to cellular and environmental cues under the control of histone chaperones and chromatin remodeling complexes, which are regulated by a wide variety of mechanisms including, but not limited to, chromatin-binding factors and enzyme complexes, covalent modifications of DNA bases, histone and nonhistone proteins, as well as coding and noncoding RNA species.[9]

Given their foundational role in chromatin organization, specific assays for nucleosome positioning can provide the most direct correlate of gene expression since their position is the integrated result of the interplay of the entire collection of DNA- and histone-modifications, and other regulatory mechanisms. A common nucleosome mapping approach involves chromatin digestion using micrococcal nuclease (MNase) to yield mononucleosomal DNA followed by next-generation sequencing (MNase-seq)[7, 11-13]. MNase is a Ca^{2+}-dependent endo-exonuclease that preferentially cleaves exposed linker DNA until it encounters a barrier such as a stably bound protein and/or a nucleosome where the histone core protects the packed DNA sequence from being digested.[14, 15] Mono-, di-, and poly-nucleosomes are generated after chromatin digestion by MNase at appropriate concentrations for a certain period. Following next-generation sequencing of libraries directed to nucleosomal DNA from digested chromatin, mapping sequenced reads of (mono-) nucleosomal DNA to a reference genome localizes protein-binding regions and nucleosome positions at up to single base pair resolution.[11] Furthermore, some form of controllable chromatin fragmentation is also critical for identifying NDRs/NFRs (e.g., by ATAC-seq[16-18] and

FAIRE-seq[19-21]) and for determining the precise genomic position of regulatory proteins, as well as histone- and DNA-modifications (e.g., by ChIP-seq[22-24]). However, macroscale nucleosome preparation, and more generally all chromatin digestion assays, have drawbacks—particularly in terms of throughput and ease of use, which can limit applicability in the analysis of challenging and often sample-limited clinical specimens. We therefore set out to design a microfluidic device to automate MNase-based chromatin processing in a way that was automated and compatible with variable cell input amounts.

Droplet microfluidics have emerged as a valuable approach for various biochemical assays, including single-cell analysis, diagnostics, DNA sequencing, and drug screening.[25-28] The basic principle involves the segmentation of an aqueous sample with an inert, immiscible oil such that each of the resulting droplets functions as an independent microreactor.[29, 30] The volume of the droplets typically range from femtoliters to nanoliters, leading to extremely small reagent consumption and amenability to rare samples, such as stem cells and circulating tumor cells.[31-37] Recent developments in the field have demonstrated exquisite control over fluid addition and removal, as well as droplet coalescence.[38-46] Additionally, rapid mixing within droplets allows more efficient chemical and biological reactions, and corresponding shorter operation times.[29, 47-49]

In this chapter, we describe the development and optimization of a droplet microfluidic device that efficiently digests chromatin to give a high yield of mononucleosomal DNA, which is the desired input material for ChIP- or MNase-seq applications. The automated device simultaneously lyses and digests whole cells to give product DNA that is indistinguishable from that obtained using benchtop processing. We also demonstrate the applicability of the approach to variable and very small cell inputs, with final validation achieved by genome-wide MNase-seq

analysis that, again, shows identical performance to bulk-processed cells. The additional validation of both PCR-free and PCR-amplified sequencing library construction makes this device of high potential value as the first step of sample preparation for a number of current and emerging epigenetic assays.

Materials and Methods

Microfluidic device fabrication

Microfluidic devices were generated using standard photolithography and soft lithography methods with details described in Chapter 2. Transparent photolithography masks were designed in AutoCAD (Schematic shown in **Figure 3-1**) and then printed by CAD/Art Services, Inc. (Bandon, OR). Negative photoresists SU8 2025 and 2050 (MicroChem Corp, Westborough, MA) were deposited onto silicon wafers (University Wafer, Boston MA) according to manufacturer recommendations. Briefly, SU8 2025 was first spin coated (PWM 32, Headway Research, Inc) onto the silicon wafer at 2000 rpm for 30 seconds to give a 40-µm thick layer. After soft baking, the resist layer was exposed to a UV lamp (Optical Associates, Incorporated) through the first-layer photomask. After post-exposure baking, the patterned wafer was developed in propylene glycol monomethyl ether acetate (PGMEA) (Sigma-Aldrich, St. Louis, MO) and dried under nitrogen. SU8 2050 was then spin coated onto the patterned wafer at 1150 rpm for 30 seconds to give a 160-µm thick layer. After soft baking, the resist was exposed to UV through the second-layer photomask and then developed as described above.

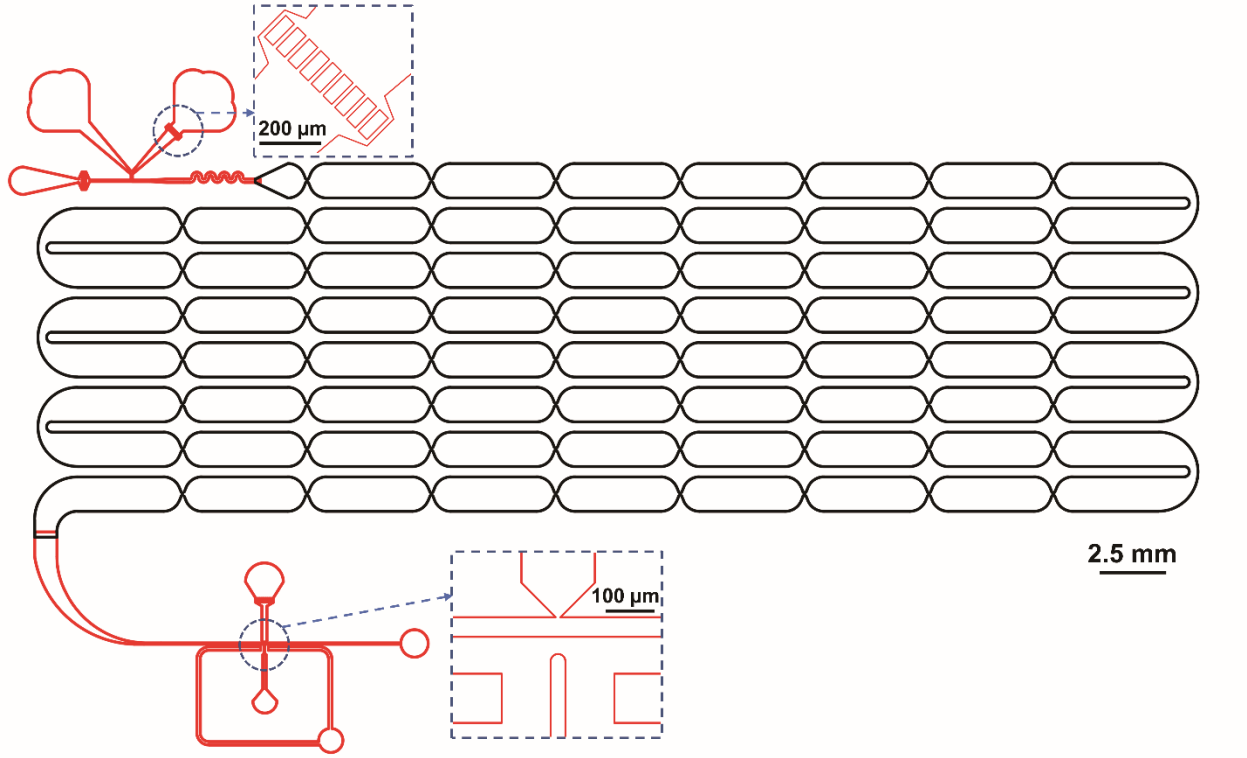

Figure 3-1 Illustration of the photomask design of the device. The red parts are the first layer, and the black part the second. The two layers are printed on two individual masks. They should be aligned to ensure the connection of the channels (shown as the overlap between the red and black). Zoom-in insertions show the design for the filter and the picoinjector.

Microfluidic devices were fabricated by casting poly(dimethylsiloxane) (PDMS) over the two-layer SU8 master. The PDMS base and elastomer curing agent (Momentive RTV 615 kit, R.S. Hughes, Carol Stream, IL) were mixed at a mass ratio of 10:1 and degassed under vacuum before the mixture was poured over the master. After curing for at least one hour at 70°C, PDMS replicas were cut and peeled from the master. Tubing inserting holes were punched with blunt-tip biopsy needles (18 gauge for inlets, 20 gauge for the outlet and electrolyte inlets). After punching holes, the devices were then sonicated in water for 5 min and dried with nitrogen to remove any PDMS debris. In parallel, a base layer was created on a glass slide by spinning a degassed, 10:1 PDMS base:elastomer mixture onto a glass microscope slide. This base layer was partially cured for 8

min before the microfluidic replica was placed on top with gentle pressure, followed by a final 30 min bake to fully bond the microfluidic device.

Microfluidic device operation

Solutions were delivered to the chip through 24-gauge PTFE tubing (Cole-Parmer, Vernon Hills, IL) controlled by syringe pumps (Pump 11 Pico Plus Elite, Harvard Apparatus). The immiscible oil phase was the fluorinated oil Novec 7500 (3M, St. Paul, MN) with 2% (w/w) perfluoropolyether polyethylene glycol block co-polymer surfactant (RAN Biotechnologies, Inc. Beverly, MA). The tubing to deliver cells to the chip was treated with 1% (w/v) Pluronic F127 (Sigma-Aldrich, St. Louis, MO)-containing PBS solution overnight and prefilled with PBS buffer. Electrolyte channels on the microfluidic device were filled with a 3M sodium chloride solution. Two clamps from electrodes of an in-house built DC-AC converter (12V DC input, 0-360V AC at 36 kHz as output), which was connected to a power supply (DC regulated power supply, TENMA T2-6628), were connected to the syringe needles supplying the electrolyte channels to form a complete circuit for the AC electric field.

Prior to experiments, microfluidic devices were treated with Aquapel (Pittsburgh Glass Works, Pittsburgh, PA) as described previously[50] to give a hydrophobic coating throughout sample processing. Optiprep (Sigma-Aldrich, St. Louis, MO) was added to PBS buffer to obtain a final density of 1.07 g/mL at a 21.9% (v/v) concentration. Cells were resuspended in this density-adjusted PBS buffer to reduce clumping and settling in the tubing. Cell suspensions were drawn into the pre-treated tubing separated from the pre-filled buffer by a plug of air. Lysis buffer (pH 7.9) contained 10 mM HEPES, 1.5 mM magnesium chloride ($MgCl_2$), 10 mM potassium chloride (KCl), and 0.5% (w/v) IGEPAL-CA630 (Sigma-Aldrich, St. Louis, MO). Digestion buffer (pH

7.5) contained 20 mM Trizma hydrochloride (Tris-HCl), 15 mM sodium chloride (NaCl), 60 mM KCl, 5 mM calcium chloride ($CaCl_2$), 0.15 mM spermine, and 0.5 mM spermidine. Quenching buffer (pH 8) contained 100 mM Tris-HCl, 20 mM ethylenediaminetetraacetic acid (EDTA), 200 mM NaCl, 2% Triton-X 100, and 0.2% sodium dodecyl sulfate (SDS). The combined lysis-digestion buffer was prepared by mixing 30 µL MNase (2000 GU/µL; New England Biolabs, Ipswich, MA), 120 µL digestion buffer, and 300 µL lysis buffer.

The fluorinated oil wets the hydrophobic channels preferentially and cuts the slower aqueous streams into droplets. Flow rates of oil (6 µL/min), cell suspension (2 µL/min), lysis-digestion reagent (2 µL/min), and quenching buffer (4 µL/min) were optimized to ensure that droplets with appropriate size were generated and quenching buffer was injected into every droplet. Droplets were collected via a 30-gauge PTFE tubing (Cole-Parmer) into a 1.5-mL tube (Eppendorf DNA LoBind, Fisher Scientific, Waltham, MA). Device operation was visualized using a high-speed camera (Phantom Miro Ex2) mounted onto either a stereoscope (Leica M80) or microscope (Leica DMi8).

Preparation of off-chip control samples

To prepare off-chip samples for comparison with those processed on-chip, an identical volume of cell suspension was added to a clean 1.5-mL tube and mixed with an equivalent volume of combined lysis/digestion buffer. Digestion was performed in bulk for the same time as droplets were incubated on-chip for enzymatic processing, and the reaction was similarly stopped by addition of quenching buffer.

Cell Culture and Sample Preparation

Jurkat acute T-cell leukemia cells (Clone E61, ATCC® TIB152, American Type Culture Collection, Manassas, VA) were cultured at 37°C with 5% CO_2 in the vendor recommended media (RPMI-1640 supplemented with 10% v/v fetal bovine serum). The cells were pelleted at 800×g for 5 min at 4°C, washed with cold PBS buffer, and resuspended into the density-adjusted PBS buffer to a defined concentration, as described below. After resuspension, cells were stored on ice until chromatin digestion. Digestion experiments were carried out with 30 µL volumes of cell suspension, unless otherwise stated.

Characterization of fragmented chromatin

After quenching and collection, droplets were coalesced by adding 50-µL 1H,1H,2H,2H-perfluoro-1-octanol (Sigma-Aldrich, St. Louis, MO). Aqueous and oil phases were separated by centrifugation at 8000 rpm for 1 min at room temperature in a tabletop microcentrifuge (Centrifuge 5418, Eppendorf, Hauppauge, NY). RNase A (10 mg/mL, Sigma-Aldrich, St. Louis, MO) and Proteinase K (10 mg/mL, ThermoFisher Scientific, Grand Island, NY) were added consecutively to the aqueous phase of both on- and off-chip processed samples to degrade RNA and proteins by incubating at 65 °C for 1 hour and 2 hours, respectively. The aqueous phases were then collected into clean tubes and DNA purified using a QIAquick PCR purification kit (Qiagen). The concentration of purified DNA was determined using the Qubit dsDNA assay (ThermoFisher Scientific, Grand Island, NY). The size distribution of the fragmented chromatin (nucleosomal DNA) was characterized using a Bioanalyzer (Agilent). All samples were diluted to or below 500 pg/µL consistent with Bioanalyzer sample submission requirements. The percentage of mononucleosomal DNA in the total DNA was obtained from the Bioanalyzer software.

Library preparation, sequencing, and bioinformatic analysis

For PCR-free library preparation, purified DNA was loaded into 1.5% agarose gel and was separated at 110V for 1 h. The mononucleosomal DNA (~147 bp) was excised from the gel and purified using the Qiagen gel extraction kit. DNA concentration was determined by the Qubit dsDNA assay, and 1 µg DNA was used for library preparation using the Illumina TruSeq® DNA PCR-Free Sample Preparation kit with a slight modification. DNA was purified using 1.8x (volumetric ratio to each library) Agencourt Ampure XP beads (Beckman Coulter, Inc.) after the end repair step, and the final library was purified using 0.55x Agencourt Ampure XP beads.

For PCR-based library preparation, 10 ng DNA was used with the NuGEN Ovation Ultralow library system V2 (San Carlos, CA). DNA was end-repaired and ligated following manufacturer instructions. After the ligation reaction, the 200-300 bp range of DNA was collected by the Agencourt RNAClean XP beads and amplified by PCR. The amplified library was purified, and the size distribution of library DNA determined by the Fragment Analyzer (Advanced Analytical Technologies, Inc.) and the concentration determined by the Qubit dsDNA assay. Four indexed libraries were pooled and sequenced 51 bp from both ends in 4 lanes of Hiseq4000 (Illumina Inc.) in the Mayo Clinic Center for Individualized Medicine Medical Genomics Facility.

Reads were aligned to the hg19 genome assembly using BWA and visualized using Integrative Genomics Browser (IGV).[51] Nucleosome peaks were identified and analyzed by Dynamic Analysis of Nucleosome Position and Occupancy by Sequencing (DANPOS) as described previously.[52,53]

Results and Discussion

Overview of the droplet microfluidic nucleosome preparation workflow

We report a droplet microfluidic device that enables automated, enzymatic chromatin processing that is amenable to variable-sized, whole cell input material, workflow shown in **Figure 3-2a**. Unfixed Jurkat cells are combined with a stream of combined lysis/digestion buffer immediately before being encapsulated into aqueous droplets at a T-junction geometry.[54] All solution flows were controlled by syringe pumps. The flow rates of cell and lysis digestion buffer were the same, and a higher oil flow rate allowed segmentation (**Figure 3-2b** and **Figure 3-3**). The tube delivering cells onto the device was pre-treated with Pluronic F-127 to reduce cells sticking to the tube.[55] A short serpentine mixing element was incorporated immediately after droplet formation to facilitate rapid mixing of cells, detergent and MNase solutions. Incubation time for cell lysis and chromatin digestion was achieved by a series of delay channels (**Figure 3-4**), with the number of delay channels increasing for longer incubation times. MNase is a Ca^{2+}-dependent enzyme, so the digestion reaction was quenched by picoinjecting[45] an EDTA buffer into passing droplets (**Figure 3-5**). A saltwater-filled electrode was used to supply an AC electric field that transiently destabilized the droplet interface to allow injection of this quenching buffer. Finally, processed droplets were collected in a 1.5-mL centrifuge tube for characterization of the resulting chromatin. The complete experimental setup, including syringe pumps and the stereoscope, is shown in **Figure 3-2c**.

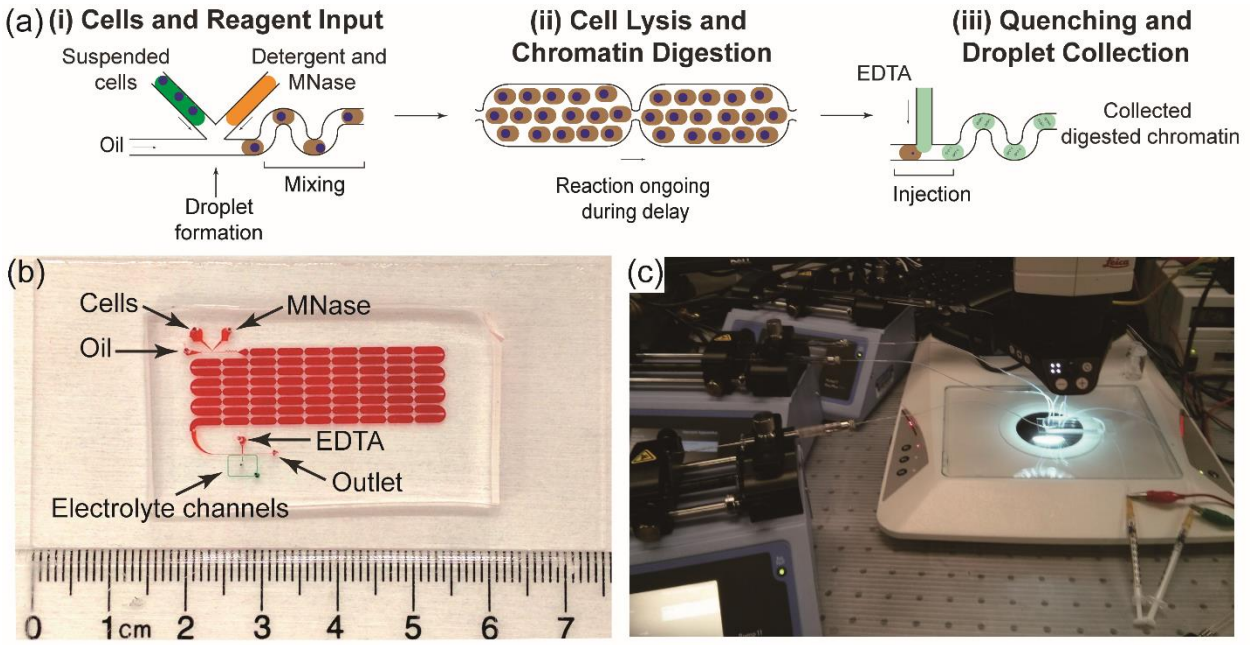

Figure 3-2 (a) Schematic of the automated, droplet-based microfluidic nucleosome preparation process. The whole procedure contains three steps: (i) loading cells, detergent, and MNase to the device for droplet generation, (ii) droplets traveling down the delay channel for chromatin digestion to complete, (iii) injecting EDTA solution to droplets to quench the enzymatic processing and collect products. (b) The 8-row device. Delay line and electrolyte channels are highlighted with red and green food dyes, respectively. (c) The complete setup for droplet microfluidic nucleosome preparation.

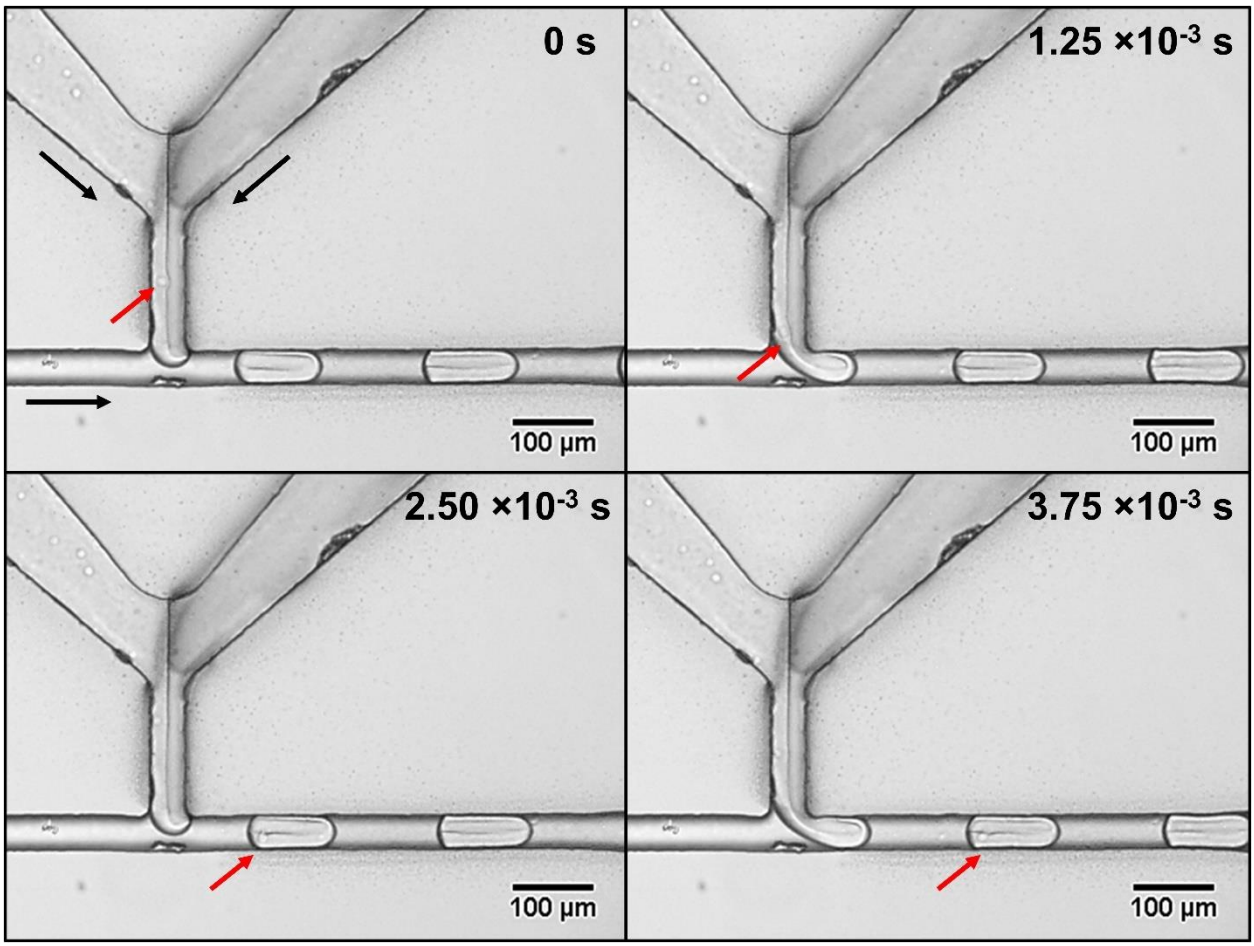

Figure 3-3 Time sequence of droplet formation in the developed enzymatic chromatin digestion device. Jurkat cell suspension contained 125,000 cells in 30-µL volume. The black arrows indicated the liquid flow directions, and the red arrows indicated the formation of one specific droplet encapsulating one cell.

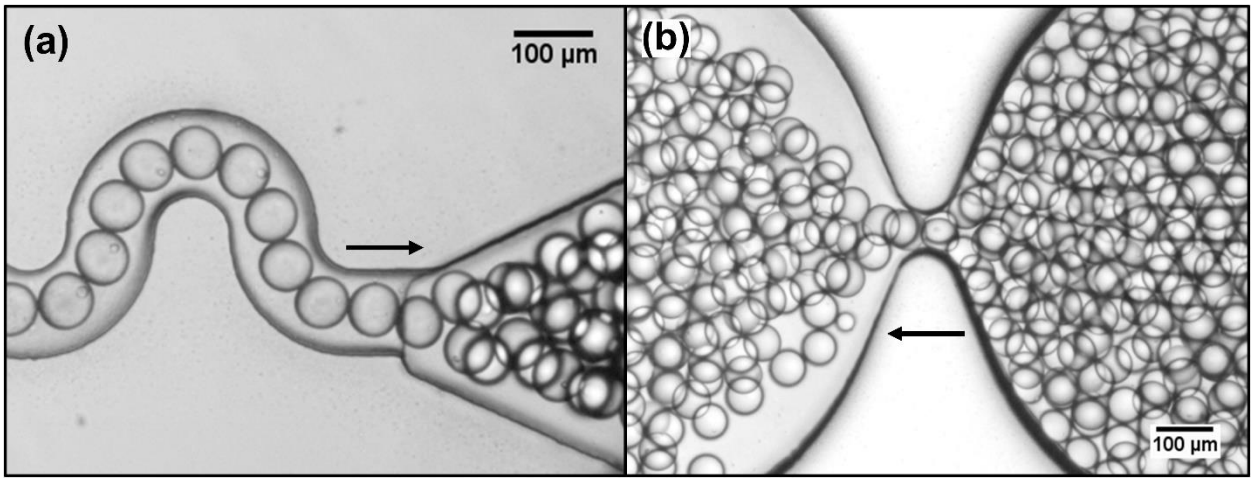

Figure 3-4 Illustration of droplets travelling across the series of delay channels. (a) Droplets entering the delay channel after generation. (b) Droplets travelling through the last constriction site to the last delay channel. The black arrows indicated the liquid flow directions.

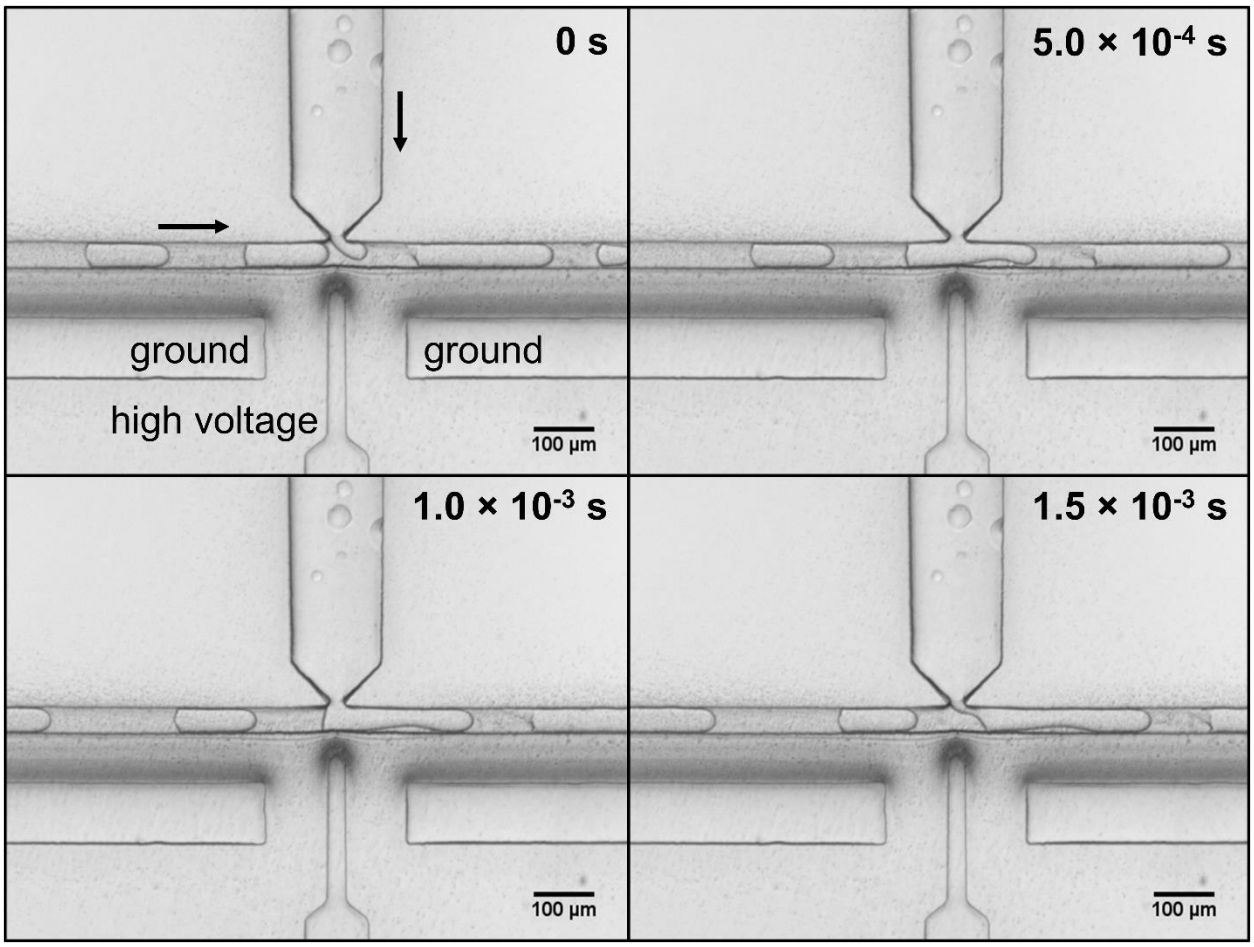

Figure 3-5 Time sequence of EDTA buffer injection to quench the enzymatic digestion reaction in passing droplets. The black arrows indicated the liquid flow directions. The electrode pairs supplied by salt water was labeld as “ground” and “high voltage”, respectively.

Multifaceted characterization of on-chip digestion efficiency

Common to standard benchtop chromatin processing is the optimization of MNase concentration and digestion time. For many applications, including nucleosome positioning assay and ChIP, DNA fragments approximating the size of DNA associated with a single nucleosome are most ideal, with ≥80% mononucleosomal DNA being an empirical benchmark.[56] To identify optimal conditions for on-chip digestion, we first titrated the concentration of MNase with a fixed delay/incubation time in the droplet microfluidic device and performed fragment analysis on purified DNA from the resulting digested chromatin. In parallel with on-chip experiments, identical digestion conditions were probed in an off-chip control experiment performed manually in Eppendorf tubes. Using a device that allowed for approximately 210 s of incubation, MNase concentrations varying from 5.3 to 266.7 GU/μL were introduced to an input of 125,000 Jurkat cells, with results shown in **Figure 3-6**. As expected, the amount of DNA fragments corresponding to the size of DNA associated with a single nucleosome (~150 bp) increased as the concentration of MNase was increased (**Figure 3-6a**). At the lowest MNase concentration, the chromatin was under-digested, with purified DNA fragment bands clearly representing the length of DNA associated with mono-, di-, and tri-nucleosomes, in addition to much larger chromatin fragments (**Figure 3-6b**, the black trace). As the concentration of MNase was increased, large fragments of chromatin were further digested to yield predominantly mononucleosomes, with >80% integrated intensity being present in the mononucleosomal DNA band at concentrations equal or greater than 133.3 GU/μL. Further increasing the MNase concentration did not result in appreciable gains in mononucleosome yield (**Figure 3-6c**). Importantly, off-chip controls showed an identical trend towards higher mononucleosome yield with increasing enzyme concentration (**Figure 3-7**), indicating that there was no bias or artifacts introduced in droplet microfluidic chromatin

processing. While not significantly different across all conditions per student t-test, the on-chip chromatin processing gave equivalent if not improved mononucleosome yield compared with off-chip controls. Closer inspection of the DNA fragment analysis for the two highest MNase concentrations showed further digestion to yield DNA sizes smaller than the typical mononucleosome. The slightly over-digested mononucleosomal DNA is desirable for the nucleosomal positioning assays described later in this manuscript but may not be optimal for other applications. However, since a 10-fold range of MNase concentration (i.e., from 26.7 to 266.7 GU/μL) generated mononucleosomes in the tested conditions, our method permits the selection of the MNase concentration most suitable for any particular application. In this study, given the desire to minimize reagent consumption, the concentration of 133.3 GU/μL was selected for further device operation.

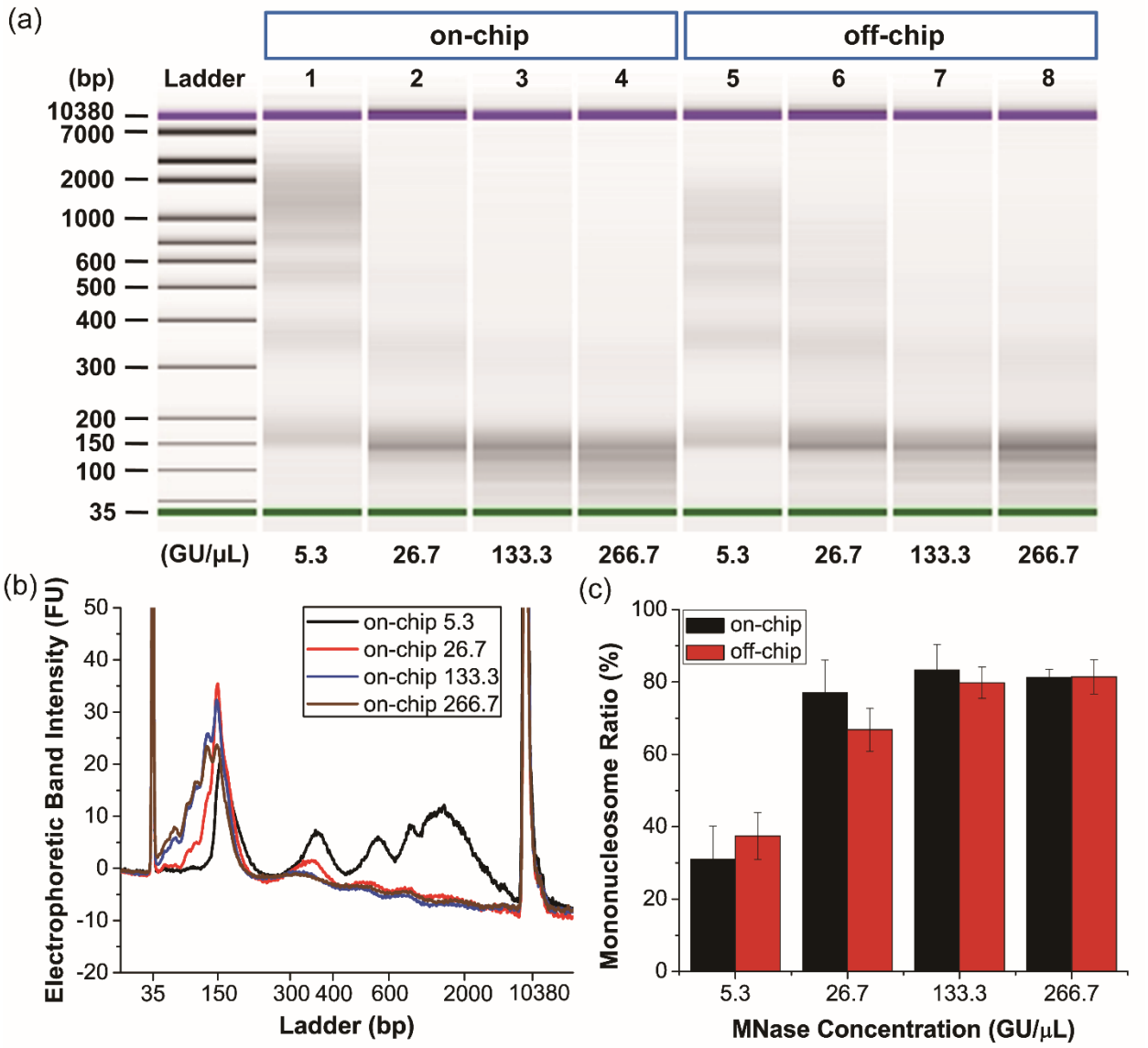

Figure 3-6 Characterization of chromatin digestion efficiency with different MNase concentrations. (a) Electropherogram of digested chromosomal DNA collected from cells processed in the droplet microfluidic devices (on-chip, lane 1-4) and 1.5-mL microcentrifuge tubes (off-chip, lane 5-8). MNase concentrations are listed on the bottom. (b) Electrophoretic band intensity profiles of on-chip samples. The peaks at 35 bp and 10380 bp are internal standards. (c) Mononucleosome yields from all samples. Error bars are standard deviation of n=3 replicates.

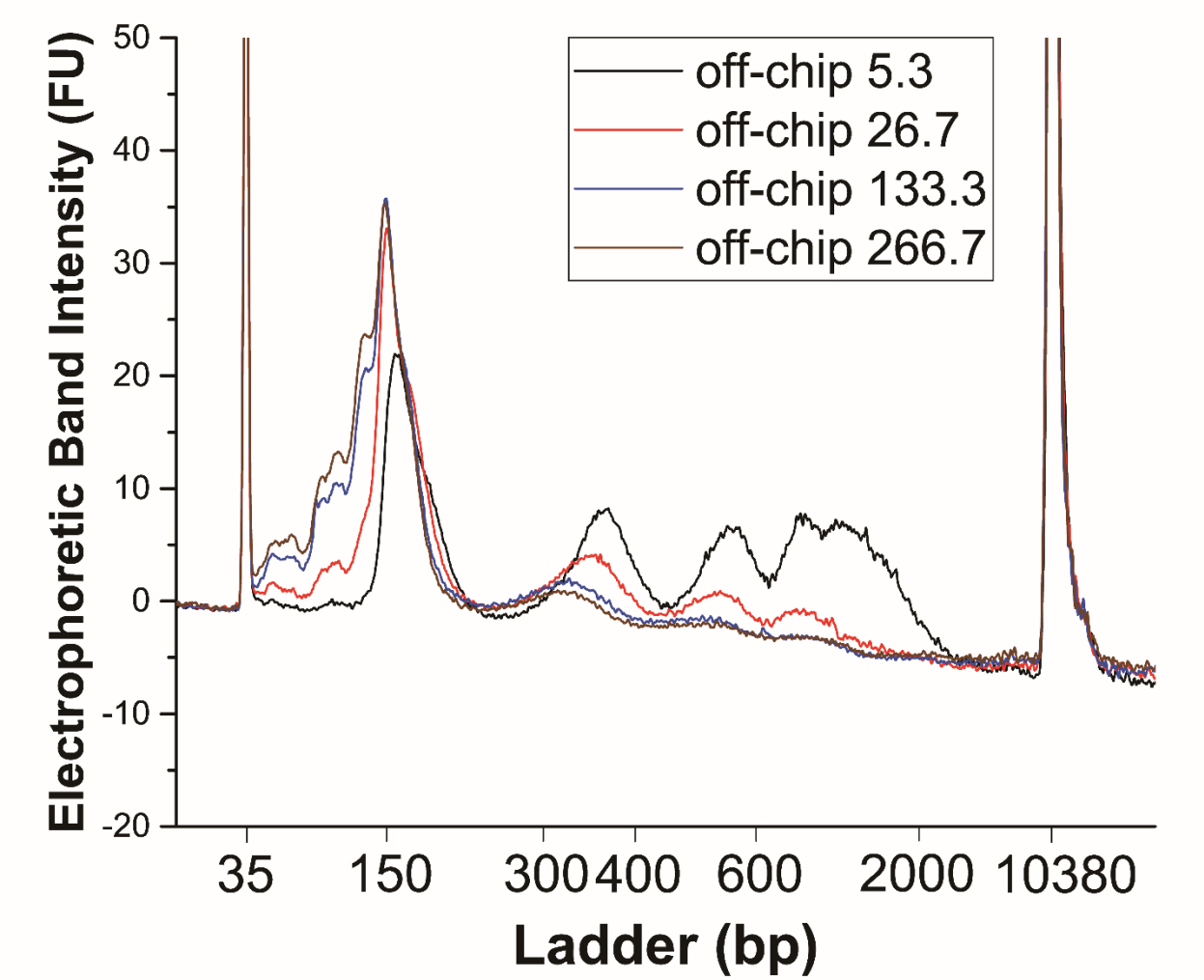

Figure 3-7 Characterization of chromatin digestion efficiency in off-chip control experiments with different MNase concentrations with a common incubation time of 3.5 minutes. Electrophoretic fragment analysis shows increasing mononucleosome yields with higher MNase concentrations. The peaks at 35 bp and 10380 bp represent internal standards.

After determining a suitable MNase concentration, we turned to optimization of incubation time. Shorter incubation periods are not only advantageous from a time perspective, but shorter delays also reduce device footprint and the potential for complications of backpressure or fabrication. Devices having four different incubation times (45, 90, 210, and 305 s) were designed and fabricated using the aforementioned protocol. A cell input of 125,000 Jurkat cells were introduced to the device and combined with 133.3 GU/µL MNase in the lysis/digestion buffer, and

DNA from the resulting chromatin was analyzed for fragment size distribution (**Figure 3-8**). At the shortest incubation time (45 s), bands for mono-, di-, and tri-nucleosomal DNA were apparent, in addition to larger fragments (**Figure 3-8a**). However, longer delay times showed increasing mononucleosome yield (**Figure 3-8b**). Again, it is important to note that the droplet microfluidic device was equivalent or superior to the off-chip control experiments (**Figure 3-8b, 9 and 10**)—particularly at shorter incubation times as the mononucleosome yield from samples processed on chip for 45 s was significantly higher from that of off chip ($p < 0.05$). This observation might be explained by the efficient mixing within droplet due to counter-rotating recirculation zones.[29] To balance the incubation time with mononucleosome yield especially the desire to obtain more slightly over-digested mononucleosomes, the device design offering a delay time of 210 s was selected for subsequent experiments.

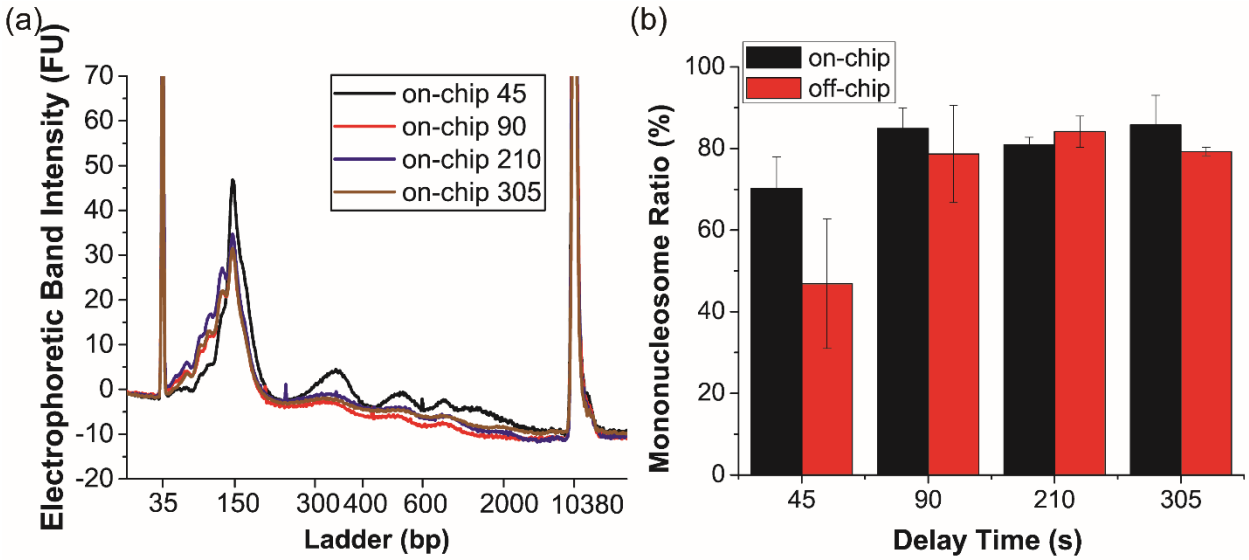

Figure 3-8 Characterization of chromatin digestion efficiency with different lengths of delay time. (a) Electrophoretic band intensity profiles of on-chip samples. The peaks at 35 bp and 10380 bp represent internal standards. (b) Mononucleosome yields from all samples. Error bars are standard deviation of n=3 replicates.

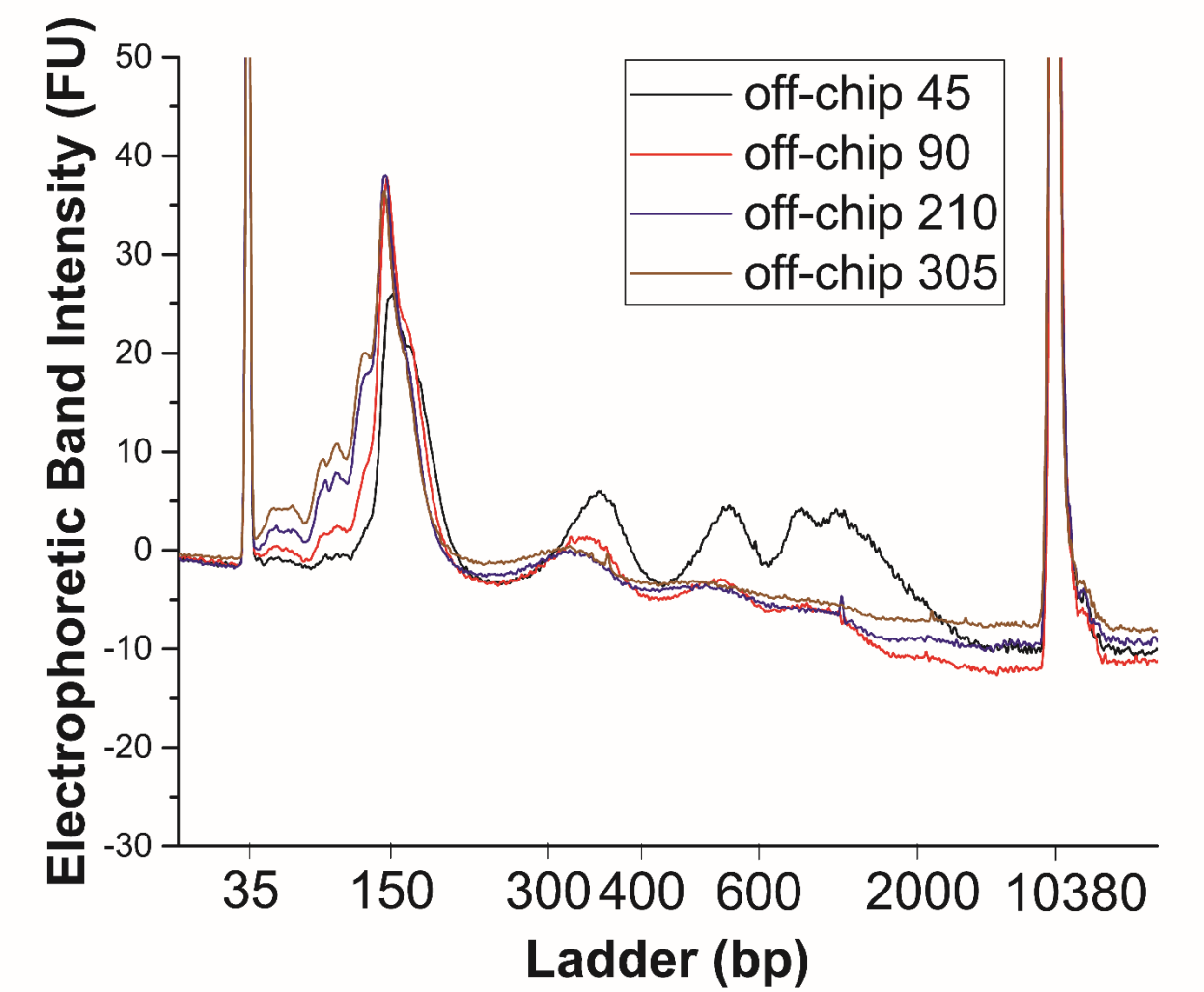

Figure 3-9 Characterization of chromatin digestion efficiency in off-chip control experiments with different incubation periods with a common MNase concentration of 133.3 GU/µL. Electrophoretic fragment analysis shows increasing mononucleosome yields at longer incubation times. The peaks at 35 bp and 10380 bp represent internal standards.

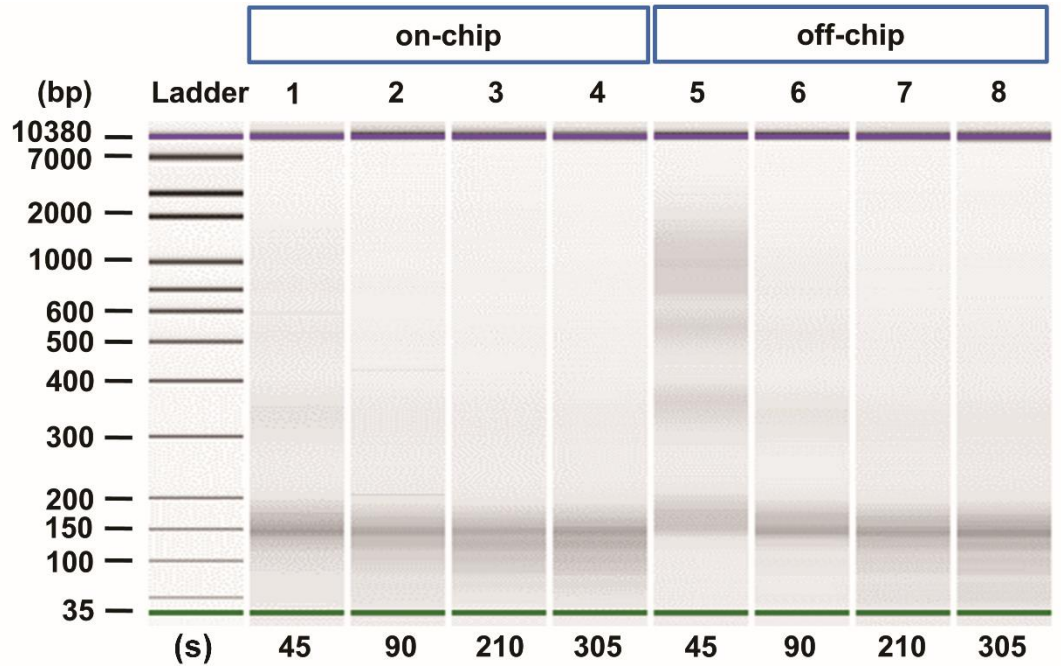

Figure 3-10 Gel image of digested chromosomal DNA collected from cells processed on chip (lane 1-4) and off chip (lane 5-8). The incubation periods are listed on the bottom. More DNA fragments of approximately or smaller than 150 bp were generated with increasing incubation time.

Validation of flexible sample input

To demonstrate flexibility in terms of cell input, we investigated the chromatin digestion efficiency using samples varying from 2,500 to 125,000 cells. Jurkat cell inputs were combined on-chip with 133.3 GU/μL of MNase and allowed to be enzymatically digested for 210 s. For consistency, these variable cell numbers were suspended in a fixed volume of density-adjusted PBS buffer. Importantly, this fixed-volume approach holds over a large input range—until a large number of droplets would have more than a single cell per Poisson statistics. Furthermore, this protocol should be amenable to clinical samples where the cell number is not precisely known. DNA fragment analysis from the resulting chromatin revealed high (~80%) mononucleosome yields across all cell inputs, with differences found not to be significant via single factor ANOVA

test ($p > 0.05$) (**Figure 3-11a**). Again, for all cell inputs, the mononucleosome yield was equivalent to or better than that observed in off-chip controls (**Figure 3-11b, 12, and 13**). Importantly, the mononucleosome yield obtained from the droplet microfluidic device is independent of cell input, and the resulting input material can be optimized for downstream applications, including ChIP and nucleosome positioning assays. It is also important to note that the minimum number of cells used in these studies is limited by the analytical characterization methods, such as the Bioanalyzer. Cells are encapsulated into droplets according to Poisson statistics and under our cell concentrations and oil and buffer flow rates cells are almost always incorporated at a single cell level (i.e. one cell per drop, with many empty drops, too). Therefore, cell digestion is almost always occurring at the single cell level and there is no reason to anticipate that this device could not be used in conjunction with single cell sequencing for single cell nucleosome positioning experiments.

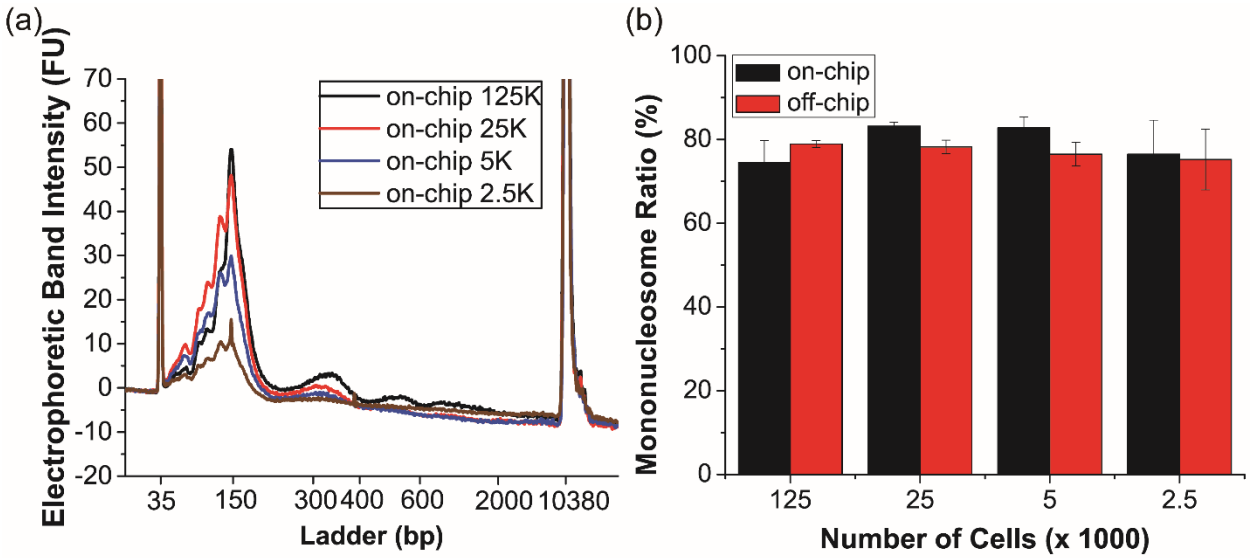

Figure 3-11 Validation of chromatin digestion with flexible cell inputs. (a) Electrophoretic band intensity profiles of on-chip samples. The peaks at 35 bp and 10380 bp are internal standards. (b) Mononucleosome yields from all samples. Error bars are standard deviation of n=3 replicates.

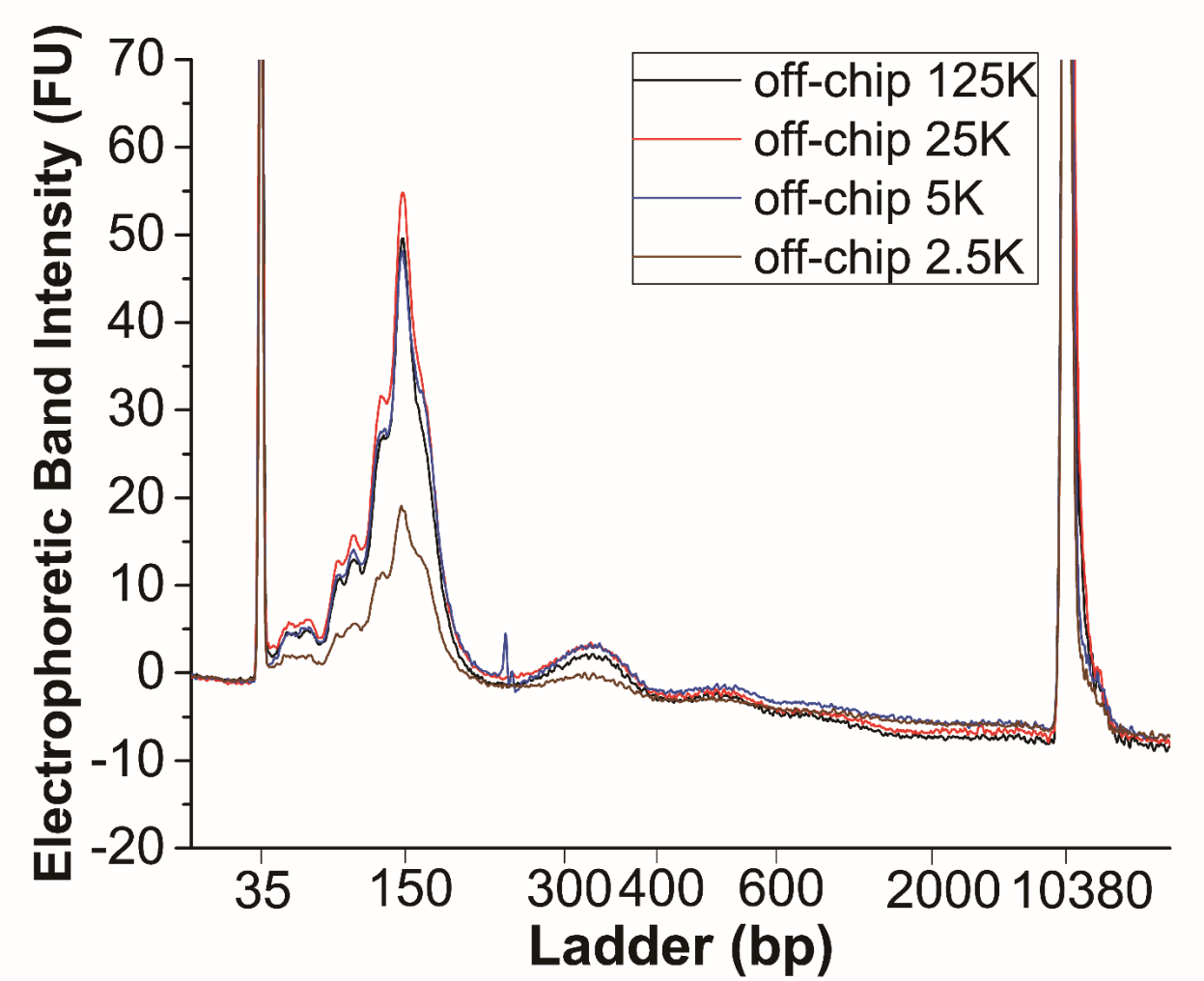

Figure 3-12 Characterization of chromatin digestion efficiency in off-chip control experiments with flexible cell inputs. The electrophoretic band intensity profiles of off chip samples confirmed the consistent efficiency in mononucleosome yields at the optimal conditions, i.e. a common MNase concentration of 133.3 GU/μL and an incubation time of 210 s. The peaks at 35 bp and 10380 bp represent internal standards.

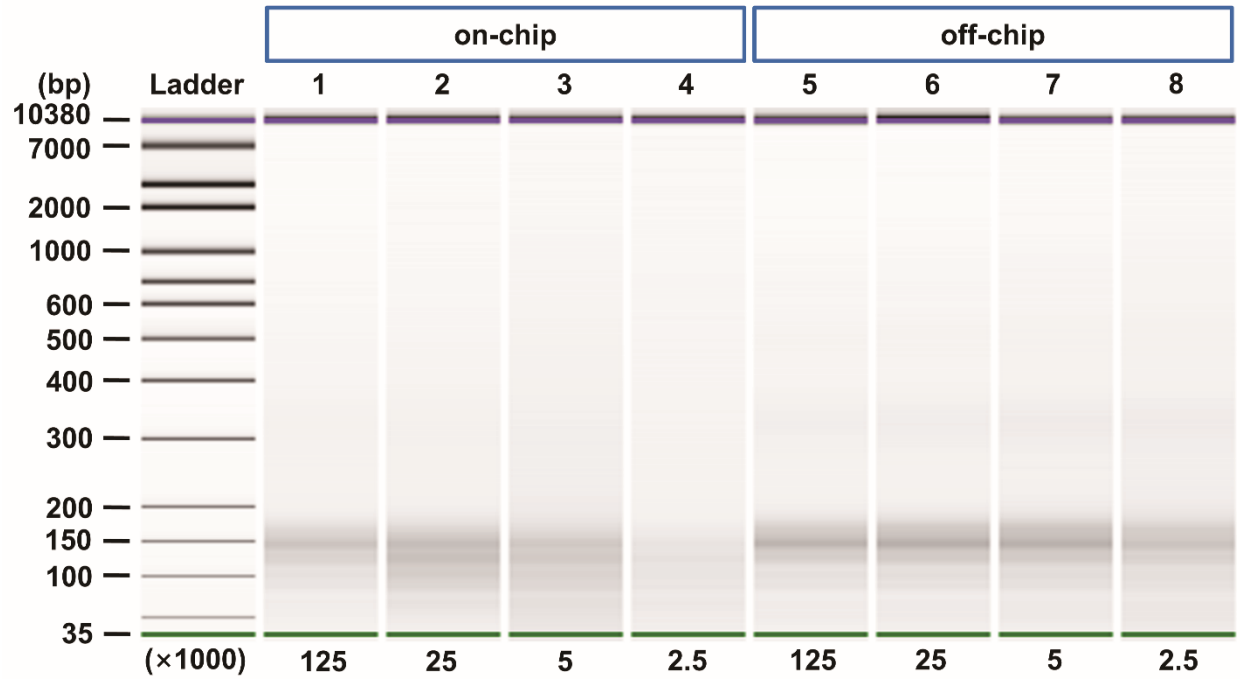

Figure 3-13 The gel image of digested chromosomal DNA with flexible cell inputs processed on chip (lane 1-4) and off chip (lane 5-8). Starting input cell numbers were listed on the bottom. Notice more DNA fragments smaller than 150 bp were presented in on-chip processed samples, suggesting overdigestion that is desired in precise mapping of nucleosome positions.

Nucleosome positioning analysis enabled by droplet microfluidically-prepared mononucleosomal DNA

Fragment analysis suggested that DNA of appropriate length and quality was produced by the droplet microfluidic MNase-digestion device; however, the ultimate validation of the resulting DNA for epigenetic analysis was further demonstrated by nucleosome positioning analysis. To ensure enough input DNA material for sequencing, the total amount of Jurkat cells for each separate digestion was increased to 500,000 (suspended in 50 μL of buffer) so that >1 μg of DNA could be collected. The digestion efficiency confirmed again this device's capability of handling flexible sized samples from perspectives of both variable volumes and cell concentrations (**Figure B-1**). Libraries for sequencing were then prepared using both PCR-free and PCR-amplified library generation protocols. The PCR-free library procedure has the advantage that it eliminates potential

biases from the PCR procedure,[57-59] whereas the PCR-amplified procedure is important to test for future applications of low input nucleosome positioning. In parallel, off-chip-prepared mononucleosomal DNA was also prepared into libraries via PCR and PCR-free. Following library preparation, samples were analyzed via next-generation sequencing on an Illumina Hi-Seq 4000 platform with one sample run per lane to obtain sufficient depth for high-resolution determination of nucleosome positioning and total nucleosome count.

PCR-free libraries resulted in 240 ± 6 million (mean ± standard deviation) uniquely mapped reads per sample (on-chip: 236 and 245 million; off-chip: 245 and 234 million). After the removal of duplicate reads, PCR-amplified libraries generated 168 ± 4 million uniquely mapped reads per sample (on-chip: 162 and 171 million; off-chip: 168 and 170 million). The resulting nucleosome positioning analyses are shown in **Figure 3-14** and **Figure 3-15**, allowing comparison of on-chip and off- chip chromatin digestion, as well as PCR-free and PCR-amplified library preparation. IGV traces of genomic regions for the constitutively expressed genes *EIF1* (eukaryotic translation initiation factor 1) and *FH* (fumarate hydratase) consistently show phased and well-positioned nucleosomes flanking the NFR associated with each gene's transcription start site (TSS) (**Figure 3-14**). This is consistent with these genes being under active transcriptional regulation. By comparison, NFRs are not observed at the TSS for the genes *GFAP* (glial fibrillary acidic protein) and *TH* (tyrosine hydroxylase), which are developmentally repressed in T cells. Further genome-wide aggregation analysis revealed remarkably well-conserved nucleosome positioning flanking the TSSs of all genes (**Figure 3-15**), with at least 5 well-positioned nucleosomes with little fuzziness flanking the TSS over a genomic region of ±1.5 kB. This observation is consistent with literature precedent using bulk-prepared nucleosomal DNA.[2, 9] Importantly, mononucleosomal DNA from both the microfluidically prepared and manually

prepared chromatin yielded highly similar numbers of sequencing reads and indistinguishable patterns of nucleosome positioning. This indicates an efficacy in using the droplet microfluidic device to generate qualified mononucleosomes for next-generation sequencing compared to traditional, manual processing.

As a final comparative metric, **Figure 3-16** shows that the number of called nucleosomes is nearly identical for on- and off-chip DNA and libraries prepared with and without PCR, with all sample preparation methods yielding approximately 15 million nucleosomes, i.e., the number expected based on the size of the human genome (approximately 3 billion bp) and the average size of nucleosomal plus linker DNA (200 bp)[60]. Furthermore, we found that the proportion of static plus partially overlapping nucleosomes in any two replicate samples (on- vs off-chip and PCR-free vs. PCR-amplified) averaged to be more than 96% (Table 3-1). Taken together, these genome-wide analyses demonstrate high reproducibility and identical performance of the droplet microfluidic device compared to manual processing in a tube. Additionally, PCR-free samples showed similar peak reads and distribution as those amplified by PCR, although PCR-amplified samples generated fewer useable reads and resulted in somewhat less well-defined nucleosome positions (**Figures 3-14 and 15**). However, the potential bias from PCR was still small, broadening the potential scope of application to smaller cell inputs. Overall, we robustly demonstrated that this droplet microfluidic nucleosome preparation strategy could provide effective samples for nucleosome positioning assays (MNase-seq) starting directly from cells, but with significant advantages in terms of automation and potential for parallel sample processing to enable higher throughput analyses.

Figure 3-14 Nucleosome profiles exemplified by four loci. Cells were digested by droplet microfluidic devices ("on") and in tubes ("off"), and the libraries were constructed with ("PCR", red) and without ("Free", green) PCR amplification. Note NFRs and the well-positioned nucleosomes flanking them near the TSSs of the consitutively expressed genes *EIF1* (positive strand) and *FH* (negative strand). NFRs are not present near the TSSs of the developmentally repressed genes *GFAP* and *TH* (both expressed from the negative strand).

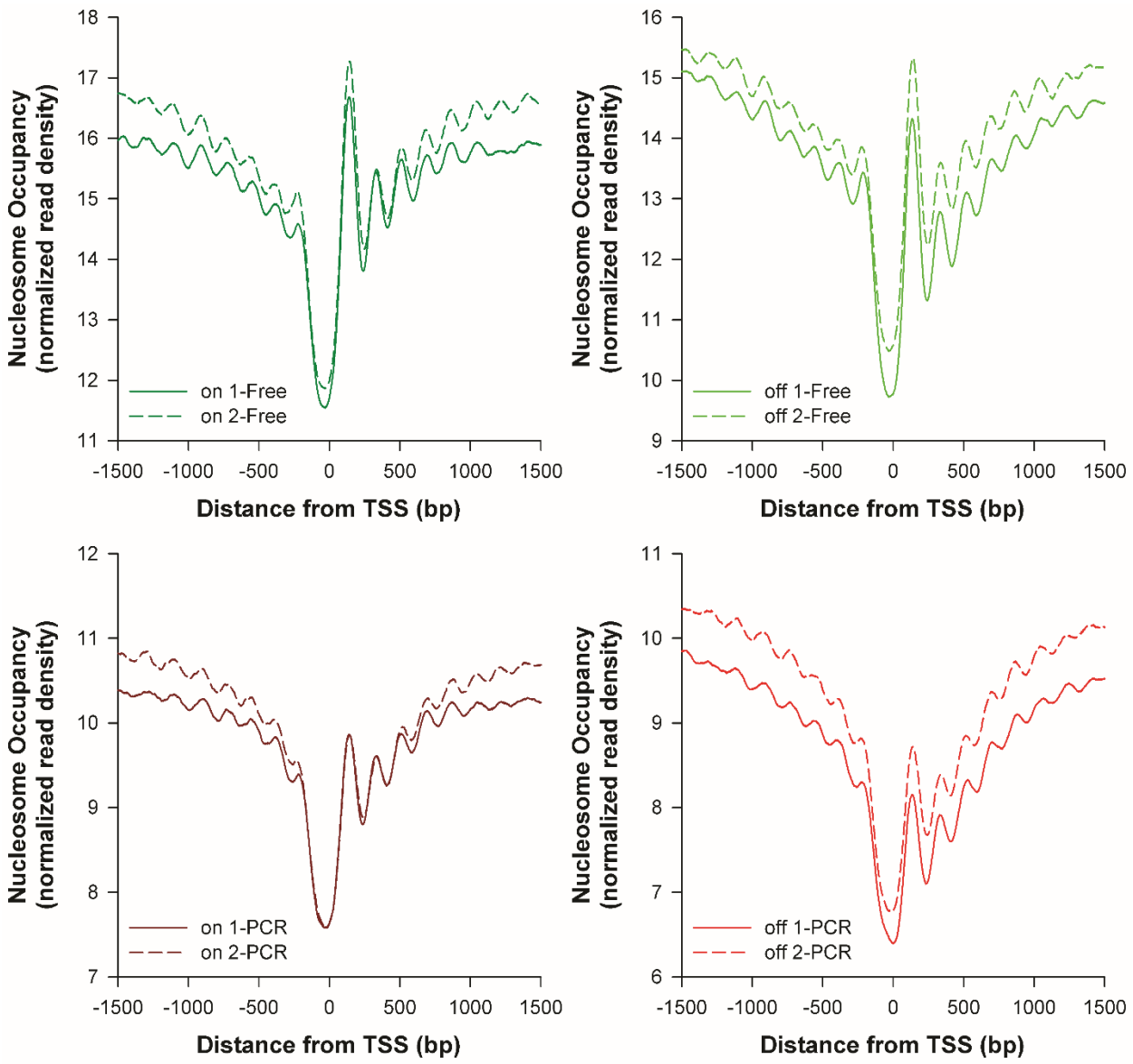

Figure 3-15 Normalized genome-wide nucleosome occupancy near TSSs in Jurkat cells processed on- and off-chip, libraries constructed with ("PCR", red) or without ("Free", green) PCR amplification.

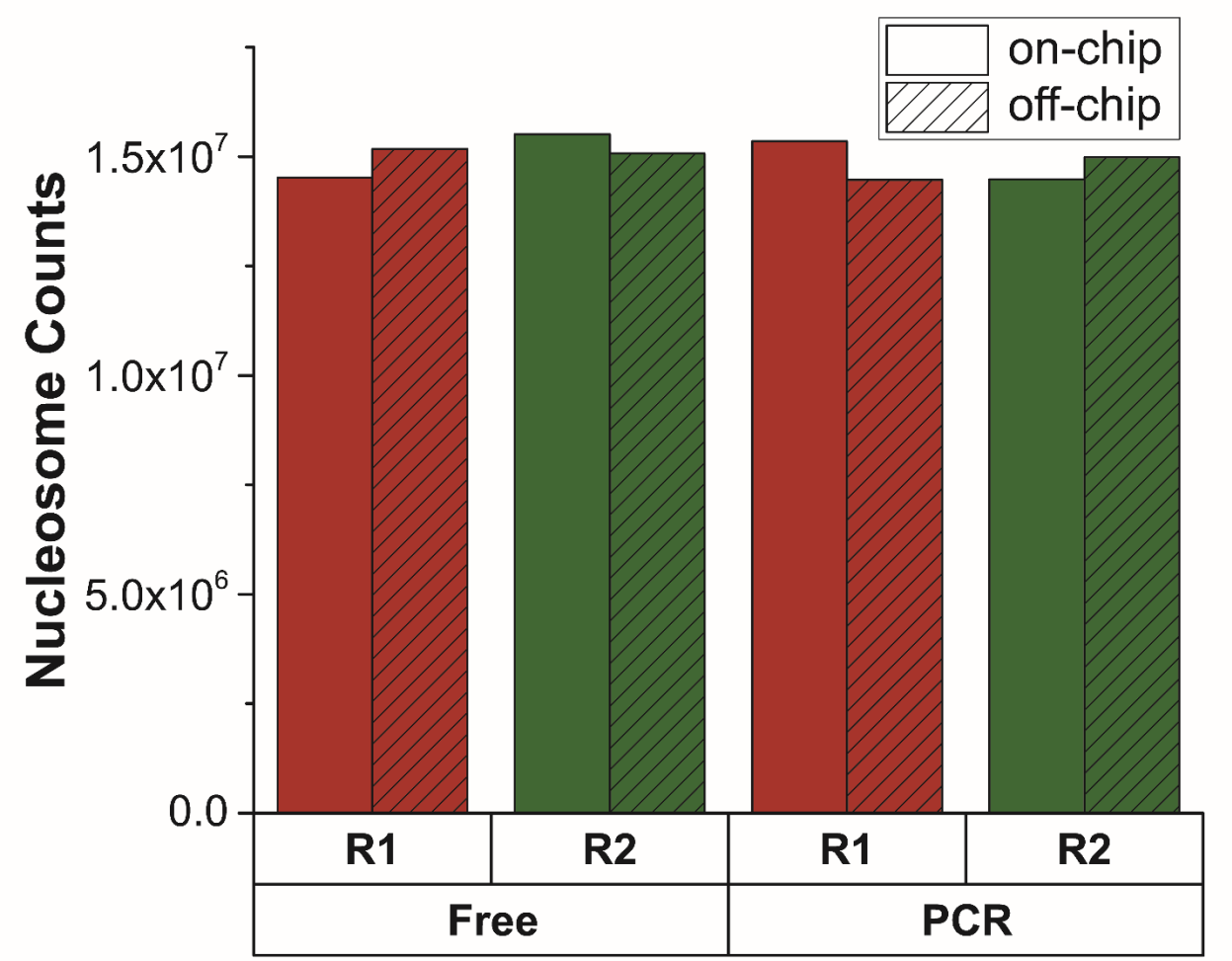

Figure 3-16 Nucleosome counts of sequenced samples that were on-chip (solid color) and off-chip (dashed) processed. Two libraries were generated from each sample (two biological replicates for on-chip and off-chip processing respectively, labeled as R1 in red and R2 in green), one directly constructed without PCR amplification ("Free") and the other with PCR amplification ("PCR").

Table 3-1 Statistical comparisons between sequenced samples.

Comparing Type	Sample A	Sample B	Overlapping Peaks*	
	ID	ID	Sample A	Sample B
Biological replicates	off 1-Free	off 2-Free	97%	98%
Biological replicates	on 1-Free	on 2-Free	97%	97%
Biological replicates	off 1-PCR	off 2-PCR	98%	99%
Biological replicates	on 1-PCR	on 2-PCR	98%	98%
off vs. on	off 1-Free	on 1-Free	97%	98%
off vs. on	off 2-Free	on 2-Free	97%	97%
off vs. on	off 1-PCR	on 1-PCR	98%	99%
off vs. on	off 2-PCR	on 2-PCR	98%	99%
Free vs PCR	off 1-Free	off 1-PCR	99%	98%
Free vs PCR	on 1-Free	on 1-PCR	98%	97%
Free vs PCR	on 2-Free	off 2-PCR	99%	96%
Free vs PCR	on 2-Free	on 2-PCR	99%	97%

* Static and partially overlapping

Conclusion

We have developed a droplet microfluidic device to enable automated, enzymatic chromatin processing starting with intact cells as input. Using Jurkat cells as a model system, we show that cells can be simultaneously lysed and chromatin digested using MNase. We optimized both enzyme concentration and delay/incubation time to maximize yield of mononucleosomal DNA above 80% to fulfill the sample requirements of many downstream epigenomic assays including nucleosome positioning assay and ChIP. We also demonstrated that the microfluidic method is insensitive to cell input, performing robustly over a range of 2,500-125,000 cells. Importantly, we found that the microfluidic device met or exceeded the performance of off-chip, bulk sample processing, as is conventional. We then utilized the collected mononucleosomal DNA

for nucleosome positioning analysis and found well-positioned nucleosomes in regions flanking TSSs of actively transcribed genes. We also found consistent nucleosome counts for on- and off-chip-processed samples and good performance for libraries prepared using both PCR-free and PCR-amplified kits. The separate success of both low input sample processing and PCR-amplified library preparation suggests amenability to future studies of nucleosome positioning from samples having limited input material. Finally, the advantages of the droplet microfluidic device in terms of automation and potential for parallelization make this a promising technology for a range of epigenetic analyses that require controllable and reproducible chromatin digestion to provide input material as the first step in downstream assays.

References

1. K. Luger, A. W. Mader, R. K. Richmond, D. F. Sargent and T. J. Richmond, *Nature*, 1997, **389**, 251-260.
2. C. Jiang and B. F. Pugh, *Nature reviews. Genetics*, 2009, **10**, 161-172.
3. K. Struhl and E. Segal, *Nature structural & molecular biology*, 2013, **20**, 267-273.
4. G. D. Bowman and M. G. Poirier, *Chem Rev*, 2015, **115**, 2274-2295.
5. C. L. Woodcock and R. P. Ghosh, *Cold Spring Harbor perspectives in biology*, 2010, **2**, a000596.
6. E. C. Small, L. Xi, J. P. Wang, J. Widom and J. D. Licht, *Proceedings of the National Academy of Sciences of the United States of America*, 2014, **111**, E2462-2471.
7. K. Maehara and Y. Ohkawa, *Scientific reports*, 2016, **6**, 19620.
8. N. Nocetti and I. Whitehouse, *Genes & development*, 2016, **30**, 660-672.
9. W. K. M. Lai and B. F. Pugh, *Nature reviews. Molecular cell biology*, 2017, **18**, 548-562.
10. J. L. Workman, *Genes & development*, 2006, **20**, 2009-2017.
11. J. G. Henikoff, J. A. Belsky, K. Krassovsky, D. M. MacAlpine and S. Henikoff, *Proceedings of the National Academy of Sciences of the United States of America*, 2011, **108**, 18318-18323.
12. C. M. Weber, J. G. Henikoff and S. Henikoff, *Nature structural & molecular biology*, 2010, **17**, 1500-1507.
13. N. A. Kent, S. Adams, A. Moorhouse and K. Paszkiewicz, *Nucleic acids research*, 2011, **39**, e26.
14. M. Noll, *Nature*, 1974, **251**, 249-251.
15. T. R. Butt, D. B. Jump and M. E. Smulson, *Proceedings of the National Academy of Sciences of the United States of America*, 1979, **76**, 1628-1632.

16. J. D. Buenrostro, P. G. Giresi, L. C. Zaba, H. Y. Chang and W. J. Greenleaf, *Nature methods*, 2013, **10**, 1213-1218.
17. J. D. Buenrostro, B. Wu, H. Y. Chang and W. J. Greenleaf, *Current protocols in molecular biology*, 2015, **109**, 21 29 21-29.
18. J. Xu, A. C. Carter, A. V. Gendrel, M. Attia, J. Loftus, W. J. Greenleaf, R. Tibshirani, E. Heard and H. Y. Chang, *Nature genetics*, 2017, DOI: 10.1038/ng.3769.
19. P. G. Giresi, J. Kim, R. M. McDaniell, V. R. Iyer and J. D. Lieb, *Genome research*, 2007, **17**, 877-885.
20. J. M. Simon, P. G. Giresi, I. J. Davis and J. D. Lieb, *Nature protocols*, 2012, **7**, 256-267.
21. Q. Zhang, T. Cheng, S. Jin, Y. Guo, Y. Wu, D. Liu, X. Xu, Y. Sun, Z. Li, H. He and Q. Xia, *Scientific reports*, 2017, **7**, 12919.
22. D. S. Johnson, A. Mortazavi, R. M. Myers and B. Wold, *Science*, 2007, **316**, 1497-1502.
23. M. F. Carey, C. L. Peterson and S. T. Smale, *Cold Spring Harbor protocols*, 2009, **4**, doi:10.1101/pdb.prot5279.
24. M. Mito, M. Kadota, K. Tanaka, Y. Furuta, K. Abe, S. Iwasaki and S. Nakagawa, *Scientific reports*, 2018, **8**, 1143.
25. S. Y. Teh, R. Lin, L. H. Hung and A. P. Lee, *Lab on a chip*, 2008, **8**, 198-220.
26. R. Seemann, M. Brinkmann, T. Pfohl and S. Herminghaus, *Rep Prog Phys*, 2012, **75**, 016601.
27. A. Huebner, S. Sharma, M. Srisa-Art, F. Hollfelder, J. B. Edel and A. J. Demello, *Lab on a chip*, 2008, **8**, 1244-1254.
28. O. J. Dressler, R. M. Maceiczyk, S. I. Chang and A. J. deMello, *Journal of biomolecular screening*, 2014, **19**, 483-496.
29. C. N. Baroud, F. Gallaire and R. Dangla, *Lab on a chip*, 2010, **10**, 2032-2045.
30. G. F. Christopher and S. L. Anna, *Journal of Physics D: Applied Physics*, 2007, **40**, R319-R336.
31. M. Leman, F. Abouakil, A. D. Griffiths and P. Tabeling, *Lab on a chip*, 2015, **15**, 753-765.
32. D. T. Chiu, R. M. Lorenz and G. D. Jeffries, *Analytical chemistry*, 2009, **81**, 5111-5118.
33. M. Chabert and J. L. Viovy, *Proceedings of the National Academy of Sciences of the United States of America*, 2008, **105**, 3191-3196.
34. H. H. Gorris and D. R. Walt, *Angew Chem Int Ed Engl*, 2010, **49**, 3880-3895.
35. E. W. Kemna, R. M. Schoeman, F. Wolbers, I. Vermes, D. A. Weitz and A. van den Berg, *Lab on a chip*, 2012, **12**, 2881-2887.
36. D. J. Eastburn, A. Sciambi and A. R. Abate, *Analytical chemistry*, 2013, **85**, 8016-8021.
37. T. M. Tran, F. Lan, C. S. Thompson and A. R. Abate, *Journal of Physics D: Applied Physics*, 2013, **46**, 114004.
38. L. Mazutis and A. D. Griffiths, *Lab on a chip*, 2012, **12**, 1800-1806.
39. V. Chokkalingam, Y. Ma, J. Thiele, W. Schalk, J. Tel and W. T. Huck, *Lab on a chip*, 2014, **14**, 2398-2402.
40. Y.-C. Tan, Y. L. Ho and A. P. Lee, *Microfluidics and Nanofluidics*, 2006, **3**, 495-499.
41. D. R. Link, S. L. Anna, D. A. Weitz and H. A. Stone, *Physical review letters*, 2004, **92**, 054503.
42. B. O'Donovan, T. Tran, A. Sciambi and A. Abate, *Journal of visualized experiments : JoVE*, 2014, DOI: 10.3791/50913.
43. L. M. Fidalgo, C. Abell and W. T. Huck, *Lab on a chip*, 2007, **7**, 984-986.
44. L. Li, J. Q. Boedicker and R. F. Ismagilov, *Analytical chemistry*, 2007, **79**, 2756-2761.

45. A. R. Abate, T. Hung, P. Mary, J. J. Agresti and D. A. Weitz, *Proceedings of the National Academy of Sciences of the United States of America*, 2010, **107**, 19163-19166.
46. S. R. Doonan and R. C. Bailey, *Analytical chemistry*, 2017, **89**, 4091-4099.
47. P. Tabeling, *Lab on a chip*, 2009, **9**, 2428-2436.
48. S. Ma, J. M. Sherwood, W. T. Huck and S. Balabani, *Lab on a chip*, 2014, **14**, 3611-3620.
49. J. D. Tice, H. Song, A. D. Lyon and R. F. Ismagilov, *Langmuir*, 2003, **19**, 9127-9133.
50. L. Mazutis, J. Gilbert, W. L. Ung, D. A. Weitz, A. D. Griffiths and J. A. Heyman, *Nature protocols*, 2013, **8**, 870-891.
51. J. T. Robinson, H. Thorvaldsdottir, W. Winckler, M. Guttman, E. S. Lander, G. Getz and J. P. Mesirov, *Nature biotechnology*, 2011, **29**, 24-26.
52. K. Chen, Y. Xi, X. Pan, Z. Li, K. Kaestner, J. Tyler, S. Dent, X. He and W. Li, *Genome research*, 2013, **23**, 341-351.
53. Z. Li, P. Gadue, K. Chen, Y. Jiao, G. Tuteja, J. Schug, W. Li and K. H. Kaestner, *Cell*, 2012, **151**, 1608-1616.
54. P. Garstecki, M. J. Fuerstman, H. A. Stone and G. M. Whitesides, *Lab on a chip*, 2006, **6**, 437-446.
55. K. Boxshall, M.-H. Wu, Z. Cui, Z. Cui, J. F. Watts and M. A. Baker, *Surface and Interface Analysis*, 2006, **38**, 198-201.
56. J. Rodriguez, J. N. McKnight and T. Tsukiyama, *Current protocols in molecular biology*, 2014, **108**, 21 28 21-16.
57. D. Aird, M. G. Ross, W. S. Chen, M. Danielsson, T. Fennell, C. Russ, D. B. Jaffe, C. Nusbaum and A. Gnirke, *Genome biology*, 2011, **12**, R18.
58. J. M. Kebschull and A. M. Zador, *Nucleic acids research*, 2015, **43**, e143.
59. D. A. Shagin, I. A. Shagina, A. R. Zaretsky, E. V. Barsova, I. V. Kelmanson, S. Lukyanov, D. M. Chudakov and M. Shugay, *Scientific reports*, 2017, **7**, 2718.
60. D. J. Gaffney, G. McVicker, A. A. Pai, Y. N. Fondufe-Mittendorf, N. Lewellen, K. Michelini, J. Widom, Y. Gilad and J. K. Pritchard, *PLoS genetics*, 2012, **8**, e1003036.

Chapter 4 A Droplet Microfluidic Platform Automating Simultaneous Sample Preparation for Two Different Epigenomic Assays

Abstract

Accessible chromatin may represent critical regulatory regions for gene expression including promoters and enhancers and coincide with nucleosome-free or nucleosome-depleted regions. Probing chromatin accessibility thus reveals the available binding sites for transcriptional machinery molecules and potential active DNA and histone modification locations nearby, therefore helps elucidate the related epigenomic regulatory network. In this chapter, a droplet microfluidic device is described in detail that can generate two samples for two epigenomic assays that provide complimentary information (MNase-based accessibility assay followed by sequencing, or MBA-seq, and MNase-seq. Samples for these two assays were referred to as MBA and MNase samples or DNA, respectively) from the same cell input sample with the same enzyme. Quality control experiments proved that the MBA samples prepared by the microfluidic device did enrich linker DNA that are related to open chromatin as the fold change of known regulatory sequences relative to reference gene region surpassed empirical threshold, and the MNase samples did get digested to mostly of mononucleosome size. Paired MBA and MNase samples were then prepared from K-562 cell lines and patient buffy coat to be sequenced and compared with existent methods for a final validation of the developed strategy for simultaneous investigation of open chromatin and nucleosome positioning.

Introduction

Eukaryotic chromatin is a packed format of nucleosomes, whose dynamic positioning and occupancy have regulatory function affecting chromatin-based biological processes including transcription, DNA repair and replication.[1] Usually, nucleosomes tend to contain histone variants and active histone marks such as H3K4me1/2 and H3K27ac to obtain destabilization and become evicted at *cis*-regulatory elements including active promoters, enhancers, and silencers.[2-4] Thus, these regulatory regions commonly coincide with nucleosome-free or depleted regions (NFR or NDR, respectively) to ensure the accessibility of DNA to non-histone proteins including the binding of transcriptional machineries and chromatin regulators.[3, 5, 6] Probing accessible DNA or nucleosome-free DNA therefore provides valuable insights in the localization of gene regulatory regions. Additionally, the crosscheck of assessible DNA regions with nucleosome positioning patterns also benefits the identification of boundary nucleosomes of open chromatin which are usually the targets of active or poised histone marks.[7] Thus, assessing accessible DNA or open chromatin plays a crucial role in elucidating genome-wide dynamics of epigenomic regulation of nucleosome occupancy and chromatin remodeling during important biological processes like gene expression, cell differentiation, and disease development.[8-12]

Common accessible chromatin investigating assays work by applying chromatin to controlled enzyme or chemical treatment to isolate the accessible DNA as these regions are more fragile to the enzymatic or chemical attack without the protection of canonical nucleosome structures. For example, DNase-seq utilizes the endonuclease DNase I to selectively digest non-occluded DNA and effectively characterizes DNase I hypersensitive sites that are historically all different types of regulatory regions.[13-16] ATAC-seq, an assay for transposase-accessible chromatin using sequencing, utilizes hyperactive Tn5 transposase to fragment and tag the genome

with sequencing adaptors simultaneously to enable fast genome-wide mapping of active regulatory elements, nucleosome positioning, and chromatin accessibility simultaneously.[17-21] FAIRE-seq, or formaldehyde-assisted isolation of regulatory elements followed by sequencing, uses formaldehyde to crosslink DNA with proteins prior to sonication-induced chromatin shearing and enrich nucleosome-depleted DNA that are of reduced crosslinking into aqueous phase by phenol-chloroform extraction before sequencing.[22] Therefore, FAIRE-seq serves as another alternative to DNase-seq and ATAC-seq without relying on any enzymatic digestion.[23]

Although these powerful epigenomic assays have revealed valuable information on open chromatin, they have their own limitations. DNase-seq can suffer from DNase I's cleavage bias leading to inaccurate transcription factor footprinting.[15] ATAC-seq may be able to map nucleosomes, but only limited to those in close proximity to accessible sites.[17] ATAC-seq also requires careful matching of cell numbers to Tn5 amount. FAIRE-seq doesn't have enzyme digestion bias, but can be affected by fixing condition and efficiency, and it has lower resolution identifying promoters compared to DNase-seq.[24] Additionally, FAIRE-seq suffers from its lower signal-to-noise ratio compared to other chromatin accessibility assays, biasing the results towards only strong signals.[25] Moreover, they all use different enzymes or chemicals from MNase, making it difficult to carry out one accessible DNA assay and the nucleosome positioning assay (MNase-seq) simultaneously on the same sample under similar conditions besides additional reagent cost. Hence, it will be beneficial and interesting to develop a method that can inspect accessible regions as well as is compatible with nucleosome positioning assay via MNase-seq.

In this chapter, sample preparation for a new assay, MNase-based accessibility assay followed by sequencing or MBA-seq that was initially developed at Mayo Clinic, Rochester, was automated using droplet microfluidics and was paired with mononucleosome preparation for

MNase-seq. The developed platform delivered two processed samples, MBA and MNase samples, at the same time with the same cell input allowing the probing of accessible DNA regions and nucleosome positioning from the same source, thus reducing data incompatibility from heterogeneity of sample replicates. By manipulating droplets in passive mode on the device, the loaded cell suspension was aliquoted into two directions, one being digested by MNase for a short period of time, the other longer. After collection of digested chromatin, sample purification and quality control by qPCR were operated to ensure the enrichment of open chromatin on the MBA side and mononucleosomal DNA on the MNase side. An optimal design of microfluidic device was selected based on quality control results for the production of paired MBA-/MNase-seq samples from K-562 cells and human buffy coat for the subsequent sequencing as a final approval of the developed droplet microfluidic platform.

Materials and Methods

Microfluidic device fabrication

Microfluidic devices were generated using standard photolithography and soft lithography methods with details described in Chapter 2 and the brief summary in Chapter 3. Photolithography masks for all tested devices were designed in AutoCAD (Schematic for the optimal device's design and naming rule were shown in **Figure 4-1**, selected representative of other designs and naming rules listed in **Figure A-2** in **Appendix A**). Briefly, SU8 2025 (MicroChem Corp, Westborough, MA) was first spin coated (PWM 32, Headway Research, Inc) onto the silicon wafer at 2000 rpm for 30 seconds to give a 40-µm thick layer. After soft baking, the resist layer was exposed to a UV lamp (Optical Associates, Incorporated) through the first-layer photomask. After post-exposure baking, the patterned wafer was developed in propylene glycol monomethyl ether acetate

(PGMEA) (Sigma-Aldrich, St. Louis, MO) and dried under nitrogen. SU8 2050 was then spin coated onto the patterned wafer at 1150 rpm for 30 seconds to give a 160-µm thick layer. After soft baking, the resist was exposed to UV through the second-layer photomask and then developed as described above. After that, the PDMS base and elastomer curing agent (Momentive RTV 615 kit, R.S. Hughes, Carol Stream, IL) were mixed at a mass ratio of 10:1 and degassed under vacuum before poured over the master. After curing and peeled off from the master, the PDMS replica was punched for tubing inserting holes with blunt-tip biopsy needles (18 gauge for inlets, 20 gauge for the outlet and electrolyte inlets). The device replicas were then sonicated in water for 5 min and dried with nitrogen to remove any PDMS debris. In parallel, a base layer was created on a glass slide by spinning a degassed, 10:1 PDMS base:elastomer mixture onto the glass microscope slide. This base layer was partially cured for 8 min before the microfluidic replica was placed on top with gentle pressure, followed by a final 30 min bake to fully bond the microfluidic device.

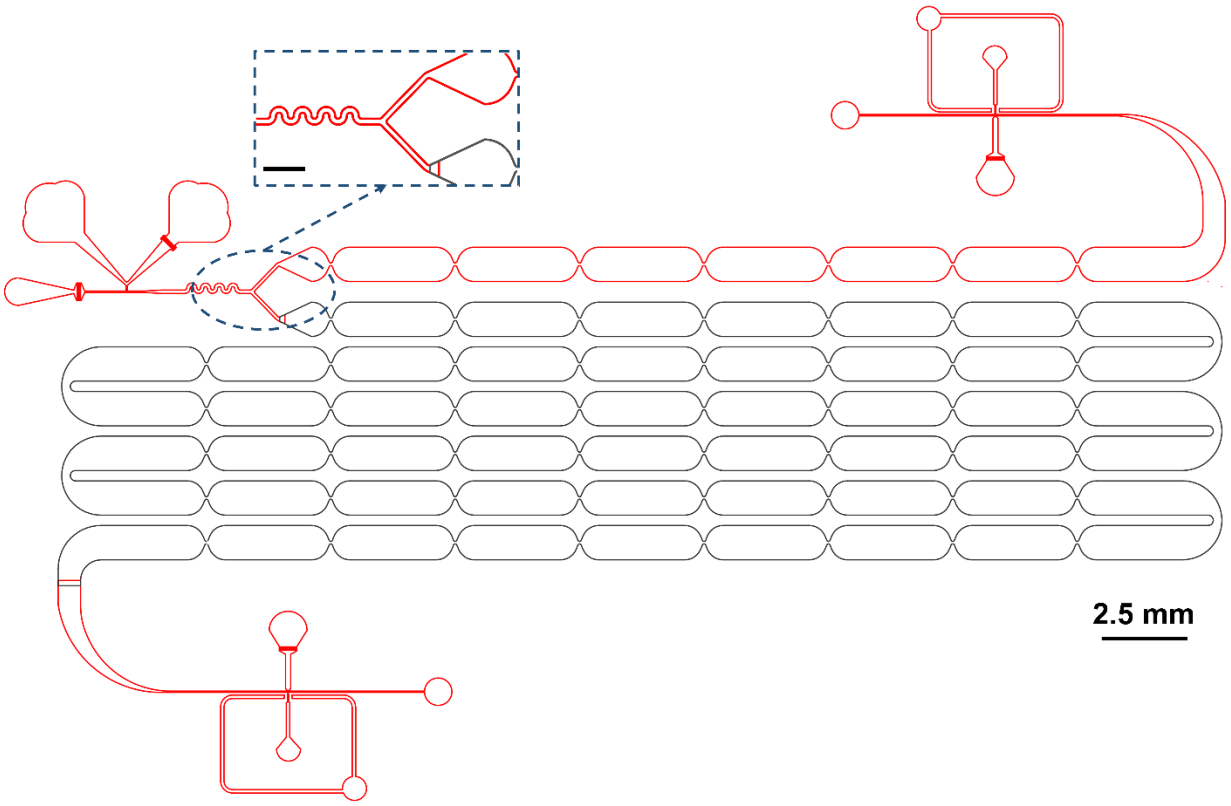

Figure 4-1 Schematic of the photomask design of the selected, optimal device "6RASD" for MBA/MNase-seq sample preparation, the name standing for "6+1 rows of delay channels, 6 rows for MNase side, 1 row for MBA side, aliquoting right after droplet generation, 'single layer' or thin, 40-μm delay for MBA side, 'dual layer' or thick, 160-μm delay for MNase side". Other designs such as "6RABD" had similar structures but with both sides being dual layered (160-μm) delay channels. The red parts were the first layer of 40-μm thickness, and the black part the second layer of 160-μm thickness. The two layers were printed on two individual masks. They should be aligned to ensure the connection of the channels (shown as the overlap between the red and black). The zoom-in insertion showed the design for droplet aliquoting to the MBA and MNase sides, respectively. The scale bar in the zoom-in insertion was 500 μm.

Microfluidic device operation

Solutions were delivered to the device through 24-gauge PTFE tubing (Cole-Parmer, Vernon Hills, IL) controlled by syringe pumps (Pump 11 Pico Plus Elite, Harvard Apparatus). The immiscible oil phase was the fluorinated oil Novec 7500 (3M, St. Paul, MN) with 2% (w/w) perfluoropolyether polyethylene glycol block co-polymer surfactant (RAN Biotechnologies, Inc. Beverly, MA). The tubing to deliver cells to the device was treated with 1% (w/v) Pluronic F127

(Sigma-Aldrich, St. Louis, MO)-containing PBS solution overnight and prefilled with PBS buffer. Two sets of electrolyte channels for the two injection sites on the microfluidic device for MBA and MNase samples respectively were all filled with a 3M sodium chloride solution. For each set, two clamps from electrodes of an in-house built DC-AC converter (12V DC input, 0-360V AC at 36 kHz as output), which was connected to a power supply (DC regulated power supply, TENMA T2-6628), were connected to the syringe needles supplying the electrolyte channels to form a complete circuit for the AC electric field.

Prior to experiments, microfluidic devices were treated with Aquapel (Pittsburgh Glass Works, Pittsburgh, PA) as described previously[26] to give a hydrophobic coating throughout sample processing. Optiprep (Sigma-Aldrich, St. Louis, MO) was added to PBS buffer to obtain a final density of 1.07 g/mL at a 21.9% (v/v) concentration. Cells were resuspended in this density-adjusted PBS buffer to reduce clumping and settling in the tubing. Cell suspensions were drawn into the pre-treated tubing separated from the pre-filled buffer by a plug of air. Lysis buffer (pH 7.9) contained 10 mM HEPES, 1.5 mM magnesium chloride ($MgCl_2$), 10 mM potassium chloride (KCl), and 0.5% (w/v) IGEPAL-CA630 (Sigma-Aldrich, St. Louis, MO). Digestion buffer (pH 7.5) contained 20 mM Trizma hydrochloride (Tris-HCl), 15 mM sodium chloride (NaCl), 60 mM KCl, 5 mM calcium chloride ($CaCl_2$), 0.15 mM spermine, and 0.5 mM spermidine. Quenching buffer (pH 8) contained 100 mM Tris-HCl, 20 mM ethylenediaminetetraacetic acid (EDTA), 200 mM NaCl, 2% Triton-X 100, and 0.2% sodium dodecyl sulfate (SDS). Elution buffer (pH 8) contained 20 mM Tris-HCl, 20 mM EDTA, 300 mM NaCl, and 2% SDS. The combined lysis-digestion buffer was prepared by mixing 30 µL MNase (2000 GU/µL; New England Biolabs, Ipswich, MA), 120 µL digestion buffer, and 300 µL lysis buffer.

Flow rates of oil (6 µL/min), cell suspension (2 µL/min), lysis-digestion reagent (2 µL/min), quenching buffer (4 µL/min), and elution buffer (4 µL/min) were optimized to ensure that droplets with appropriate size were generated, droplets travelled across the delay channels at both directions for appropriate time to allow limited or thorough MNase digestion respectively, and quenching buffer and elution buffer were injected into every droplet at both MBA and MNase injection sites. Droplets were collected via two 30-gauge PTFE tubing (Cole-Parmer) into two 1.5-mL tubes (Eppendorf DNA LoBind, Fisher Scientific, Waltham, MA) respectively. Device operation was visualized using a high-speed camera (Phantom Miro Ex2) mounted onto either a stereoscope (Leica M80) or microscope (Leica DMi8).

Cell culture, buffy coat extraction, and microfluidic sample preparation

K-562 chronic myelogenous leukemia lymphoblast cells (ATCC® CCL-243™, American Type Culture Collection, Manassas, VA) were cultured at 37°C with 5% CO_2 in the vendor recommended media (ATCC-formulated Iscove's Modified Dulbecco's Medium, Catalog No. 30-2005, supplemented with 10% v/v fetal bovine serum). The cells suspended in 10 mL medium were fixed with 625-µL 16% formaldehyde (Methanol-free, Pierce™, Catalog No. 28906, ThermoFisher Scientific) for 5 min at room temperature. Fixation reaction was quenched by adding 500-µL 2.5M glycine solution and incubated for 5 min at room temperature. The cells were pelleted at 800×g for 5 min at 4°C, washed with cold PBS buffer, and resuspended into the density-adjusted PBS buffer to a defined concentration, as described below. After resuspension, cells were stored on ice until chromatin digestion. Digestion experiments of K-562 cells were carried out with 50 to 100 µL volumes of cell suspension at a density of 4 to 5 million cells per mL, unless otherwise stated.

Buffy coat (white blood cells and platelets, the blood components other than red blood cells or plasma) was extracted from whole blood samples of healthy donators via venipuncture and collected in anti-coagulant containing blood collection tubes (K2EDTA tubes with purple caps) by the Department of Emergency Medicine of the University of Michigan Taubman Center and Michigan Center for Integrative Research in Critical Care (MCIRCC). Blood samples were then aspirated into a 50-mL conical tube and treated with ACK lysing buffer (Catalog No. A1049201, Thermo Fisher Scientific) based on the manufacturer's recommended protocol. Briefly, 1-mL whole blood was mixed with 12-mL ACK lysing buffer and incubated at room temperature for 5 minutes. Buffy coat was collected by centrifugation at 300 x g for 5 minutes at room temperature and the removal of supernatant. Additional 3-mL ACK lysing buffer could be added to the buffy coat for another 2-minute incubation if significant numbers of red blood cells remained on the top. After centrifugation and supernatant aspiration, buffy coat cells washed with 5-mL cold PBS buffer and centrifuged for the pellet. The washed buffy coat was then resuspended in a specific volume of density-adjusted PBS buffer (see Chapter 3 for detailed recipe) for desired cell concentration. Digestion experiments of buffy coat were carried out with 60 to 100 µL volumes of cell suspension at a density of 10 million cells per mL, unless otherwise stated. The usage of human subject materials for this research was approved by the Health Sciences and Behavioral Sciences Institutional Review Board (IRB-HSBS) on the University of Michigan – Ann Arbor Campus.

Characterization of fragmented chromatin

Collected droplets from both directions were coalesced by adding 50-µL 1H,1H,2H,2H-perfluoro-1-octanol (Sigma-Aldrich, St. Louis, MO). Aqueous and oil phases were separated by centrifugation at 8000 rpm for 1 min at room temperature in a tabletop microcentrifuge (Centrifuge

5418, Eppendorf, Hauppauge, NY). RNase A (10 mg/mL, Sigma-Aldrich, St. Louis, MO) was added to the aqueous phases of both MBA and MNase samples to degrade RNA by incubating at 65 °C for 1 hour. Proteinase K (10 mg/mL, ThermoFisher Scientific, Grand Island, NY) was added consecutively to the aqueous phases of both MBA and MNase samples to degrade proteins by incubating at 65 °C for 2 hours. The samples were maintained on the hotplate at 65 °C for another hour to ensure full reverse-crosslinking of DNA and protein residues. The aqueous phases were then collected into clean tubes. DNA for MNase-seq were purified using a QIAquick PCR purification kit (Qiagen), and DNA for MBA-seq were purified using a QIAquick Nucleotide Removal kit (Qiagen). The concentration of purified DNA was determined using the Qubit dsDNA assay (ThermoFisher Scientific, Grand Island, NY). The size distribution of the DNA purified from fragmented chromatin (both limited digested DNA containing nucleosome-free regions for MBA-seq, and mononucleosomal DNA for MNase-seq) was characterized using a Bioanalyzer (Agilent). All samples were diluted to or below 500 pg/µL consistent with Bioanalyzer sample submission requirements. The percentage of mononucleosomal DNA in the total DNA was obtained from the Bioanalyzer software.

Size selection of linker DNA

QIAquick purified DNA from MBA samples were first incubated with 2.5× volumes (compared to DNA solution volume) of room-temperature Agencourt Ampure XP beads (Beckman Coulter, Inc.) for 15 minutes to allow high bp DNA to bind with the beads (approximately > 100 bp). The linker DNA was collected by centrifuging the beads and using a magnet to attract the beads for 5 min and taking the suparnant. The volume of the supernatant was made up to 400 µL by adding EB buffer from QIAgen PCR purification kit (notice that this is EB

buffer, not the in-house made elution buffer described in section "Microfluidic device operation"), and the solution was mixed with 6-µL glycogen (Catalog No. R0561, ThermoFisher Scientific), 8-µL 5M NaCl, and 800-µL 100% cold ethanol freshly taken from -20 °C freezer and incubated overnight. The pellet of glycogen containing linker DNA was collected by centrifugation of the mixture at maximum spin rate for 13 minutes at 4 °C and pouring out the supernatant. After washing the pellet with 700-µL 70% ethanol taken from -20 °C freezer and centrifuging at max spin rate for 5 minutes at 4 °C, the solution was pipetted out carefully to leave the pellet, which was dried for 5 minutes at room temperature by opening the tube lid and exposing to air. The first-round size selected linker DNA was eluted with 10.5-µL EB buffer from QIAgen by dissolving the glycogen pellet into the EB buffer, incubating for 5 minutes, spinning briefly, and pipetting out 10-µL aqueous phase into a clean DNA Lobind tube after putting the solution onto a magnetic stand to avoid leftover beads' contamination.

The purified DNA was then constructed into Illumina compatible libraries with ThruPLEX DNA-seq kit (Rubicon Genomics) based on the manufacturer's recommended protocol with slight modifications: in the third section of "Library Amplification Reaction", the last step of "Addition of Indexes" was shortened to 20 s, and the last step in "Library Amplification" was shortened to 5-10 s to selectively amplify the 30-100 bp linker DNA. Libraries were then purified with Ampure XP beads following Rubicon's recommended protocol. After collecting the 100-300 bp bands of the 2% agarose gel of the libraries, the adaptor ligated linker DNA were extracted with QIAgen's MinElute gel extraction kit (Qiagen) and eluted to 15-µL EB buffer. These linker DNA for MBA-seq were ready for quality control.

Quality control of collected MBA and MNase DNA

Two known accessible loci (CD47 enhancer and DZIP3) were used as positive targets to assess the enrichment of nucleosome-free, accessible regions of MBA samples. An inaccessible locus (C19 intergenic region[27]) was used as reference evaluating the digested chromatin's background. Sequences of primer pairs for these three regions were listed in **Table B-1** in **APPENDIX B**.

Quantitative PCR reaction per well on the 384-well reaction plate (Applied Biosystems Catalog No. 4309848, Fisher Scientific) was set up according to the manufacturer's recommended parameters. Briefly, for each target, 10-µL PowerUp™ SYBR™ Green master mix (Applied Biosystems Catalog No. A25776, Fisher Scientific) was mixed with 2-µL forward primer (5 µM) and 2-µL reverse primer (5 µM). Then 2 to 3 µL of DNA template was added to the mix depending on the DNA concentration to ensure appropriate amount of DNA was used. The volume was made up to 20 µL by adding appropriate volume of nuclease-free water. After setting up the PCR plate, the experiment was run on 7900HT Fast Real-Time PCR system (Applied Biosystems) with the standard parameter setup at the DNA sequencing core facility (University of Michigan). Fold change of positive loci (posi) between MBA DNA (BA) and MNase DNA (MN) was calculated using equation (**4.1**) to assess the efficiency of accessible region enrichment in the MBA DNA.

$$log_2\ fold\ change = c_{t(BA,reference)} - c_{t(BA,posi)} - (c_{t(MN,reference)} - c_{t(MN,posi)}) \quad (4.1)$$

Results and Discussion

Overview of the device operation

The optimal droplet microfluidic device shown in Figure 4-1 automated sample processing for MBA-seq and MNase-seq from cell input directly allowing simultaneous investigation of nucleosome-free and -associated regions of the same sample for regulatory region localization and

nucleosome positioning mapping, respectively. Fixed K-562 cells were loaded to the device through the left inlet to be coflowed with lysis-digestion reagent containing MNase and encapsulated into droplets cutting by the oil flow (**Figure 4-2a**). The droplets were directed to two different delay channels at the intersection (**Figure 4-2b**), the upper one for MBA-seq sample preparation as the thinner and shorter (40-μm, 1 row) delay channel only allowed < 10 s of reaction time for MNase to attack the most fragile regions of cellular chromatin, i.e., accessible regions that were nucleosome-free, while the lower one for MNase-seq sample preparation as the thicker and longer delay channel allowed sufficient time (> 5 min) for MNase to thoroughly digest chromatin and produce (mono-)nucleosomal regions that were protected. At both directions picoinjectors were designed to inject elution buffer or quenching buffer to the droplets (**Figure 4-2c and d**). Elution buffer contained higher concentration of SDS, thus was supposed to completely stop any enzymatic activity of MNase to preserve the fragile accessible regions, which was significant in their enrichment.

Figure 4-3 showed the droplets travelling in the delay channels with intact integrity. Depending on the difference in length and thickness between MBA and MNase delay channels as well as droplet size and the location of aliquoting site, the ratio of numbers of droplets aliquoted between the two directions and the corresponding delay times would be varied. For the optimal design shown here, approximately 60% of produced droplets were directed to the MBA side. Notice that due to the existence of empty droplets and the different digestion levels between MBA and MNase samples, the ratio of numbers of droplets aliquoted between the two assays did not necessarily correlate with the ratio of the amount of DNA quantified for the two assays.

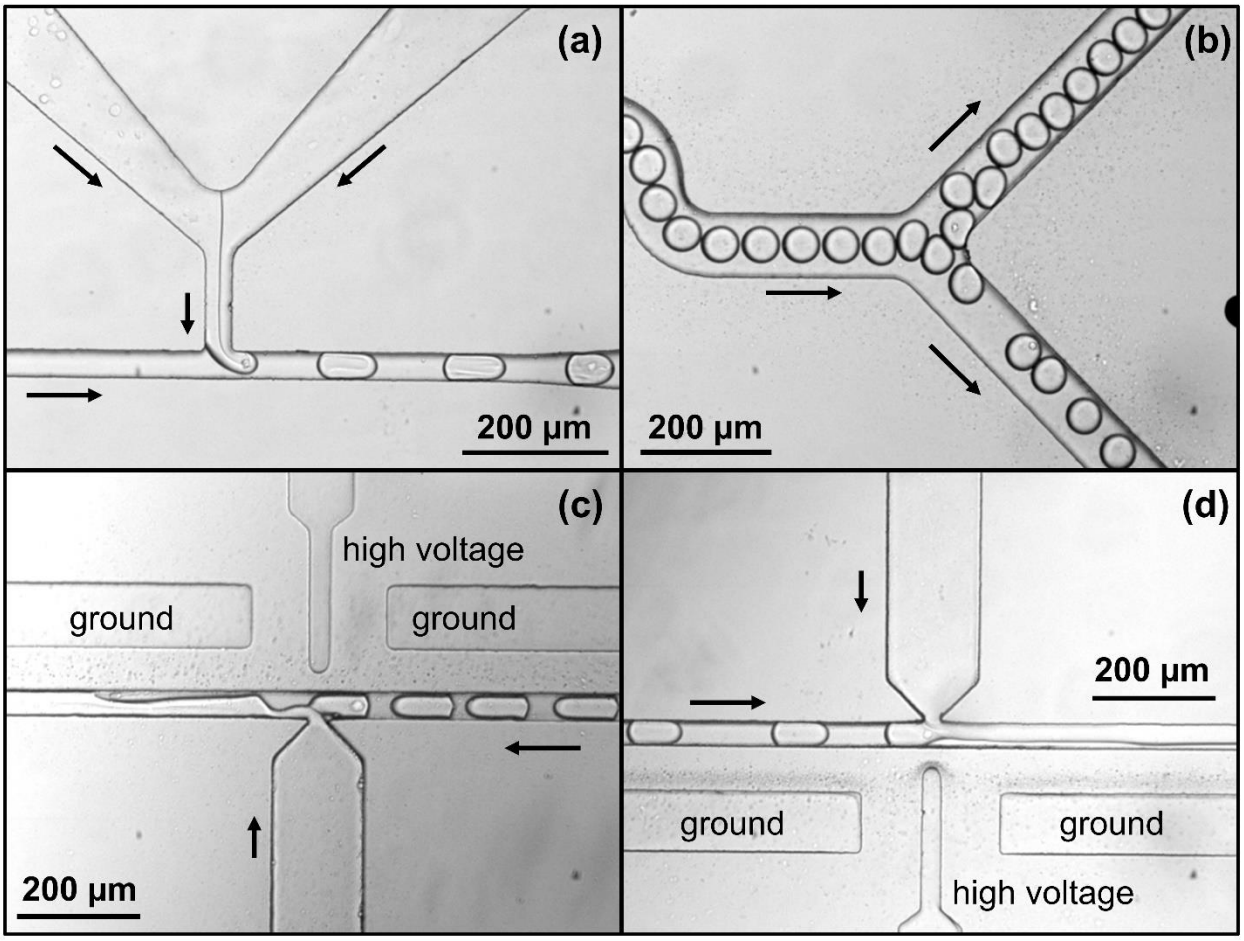

Figure 4-2 The operation of key components of the MBA/MNase sample preparing device. (a) K-562 cells (4 million per mL) encapsulated into droplets. (b) Droplets aliquoted into the two directions for MBA and MNase sample processing. The upper branch directed droplets to the (c) MBA injection site as the delay time was only < 10 seconds in the one-row, single-layered delay channel, while the lower branch directed droplets to the (d) MNase injection site as the delay time was > 5 minutes in the six-row, thicker delay channel. The black arrows indicated the flow directions of the liquid.

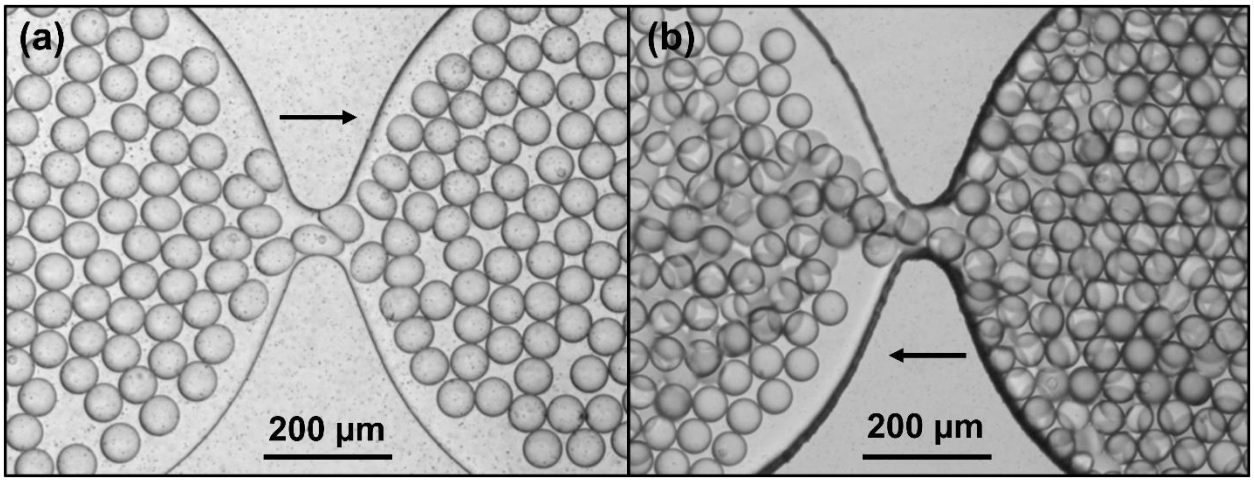

Figure 4-3 (a) Droplets travelling across the third constriction site of the delay channel on the MBA side. (b) Droplets travelling across the last constriction site of the delay channel on the MNase side. The black arrows indicated the flow direction of the oil and the droplets.

Chromatin Digestion Efficiency of Devices with Different Designs

As being described above, the accessible regions are more fragile to MNase attack compared to nucleosomal regions, thus the delay time for MBA sample side needed to be controlled carefully to preserve as much nucleosome-free regions as possible to avoid targeted sample loss or region bias in the subsequent analysis. Therefore, multiple designs of droplet microfluidic devices were tested to ensure that both accessible regions and mononucleosomes were achieved from the on-chip processed samples. Two criteria were applied here to assess the quality of both DNA products, one is digestion level, the other is quantitative PCR on known accessible and nucleosomal regions.

In terms of digestion level, MBA DNA is expected to have limited digestion, empirically with little (~ or < 20%) to no visible peak at mononucleosomal range (about < 250 bp, mononucleosome length with linker DNA length), while MNase DNA is expected to have mononucleosome as the main product. **Figure 4-4** showed the box plots with data points of digestion levels in the form of mononucleosome yield of samples produced from different droplet

microfluidic devices. MBA and MNase DNA were generated in pairs from the same device, i.e., "4R_BA" and "4R_MN" contained two pairs of MBA and MNase DNA generated from two cell input replicates loaded to two devices of the same design, respectively. Each pair of data points stood for one replicate. Notice that "6RASD_BA" had two data points "missing" compared to "6RASD_MN", because those two data points did not have effective values as the mononucleosomal region contained little area and could not be extracted effectively from the Bioanalyzer electropherogram profiles, indicating very limited digestion which could be beneficial. Moreover, considering the less MNase digestion efficiency for nucleosome positioning assay of fixed cells compared to native cells empirically (see **Chapter 3**), ~ 70% mononucleosome yield could be qualified. Additionally, MNase digestion efficiency should also be balanced with device footprint, therefore the length of MNase side delay channel should not be unlimitedly elongated. With these factors taken into consideration, as well as experimental reproducibility and data outliers revealed by Figure 4-4, 6RABD and 6RASD had the most potential to generate qualified samples for both assays.

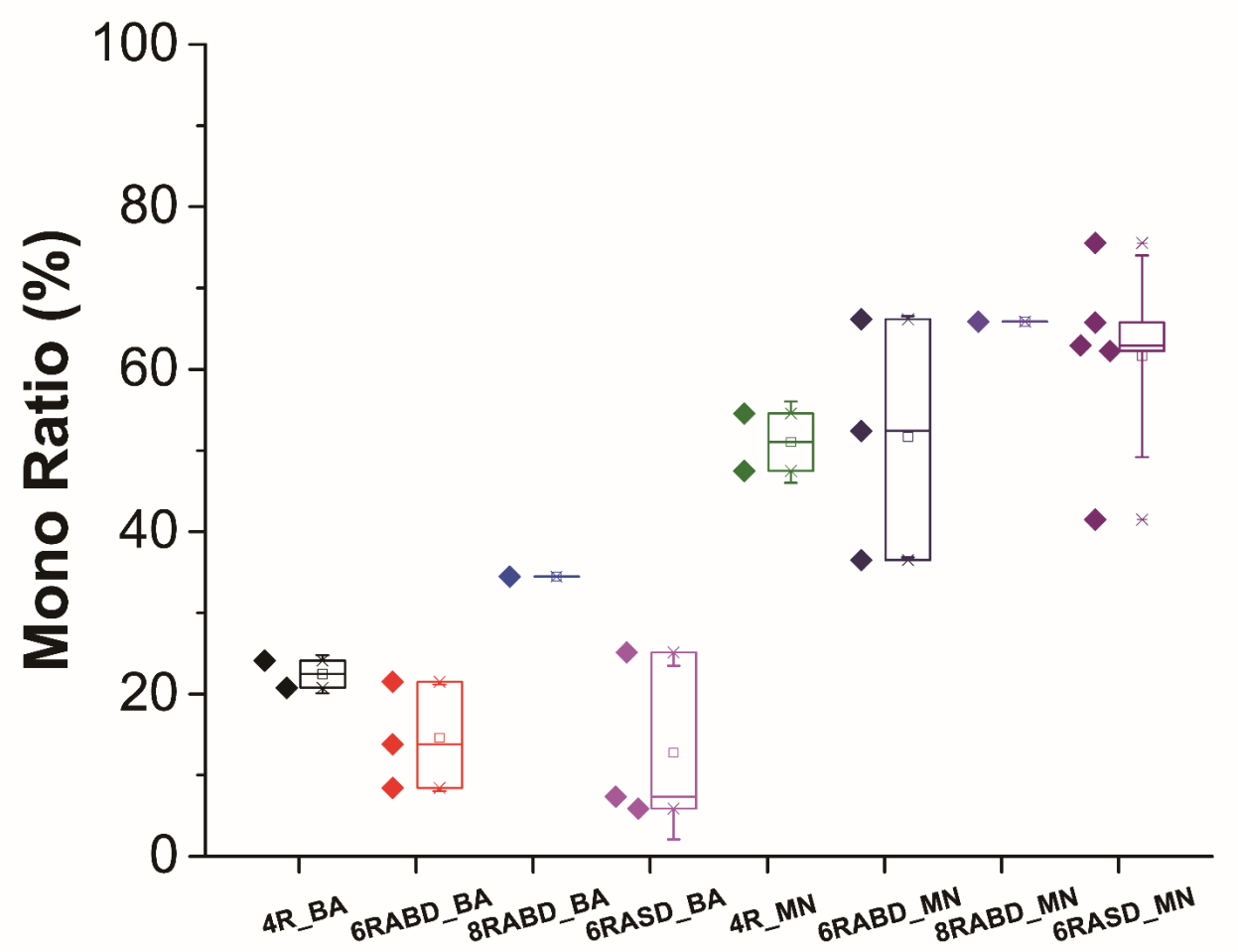

Figure 4-4 Mononucleosome yield of MBA-seq ("BA") and MNase-seq ("MN") samples from different devices. "4R" had 4 rows of thick, 160-µm delay channels in total, aliquoting at the end of the first row. "6RABD" had 6+1 rows of delay channels, 6 rows for MNase side, 1 row for MBA side, aliquoting right after droplet generation, both sides contained "dual layer" or thick delays. "8RABD" had 8+1 rows of delay channels, 8 rows for MNase side, 1 row for MBA side, aliquoting right after droplet generation, both sides contained "dual layer" or thick delays. "6RASD" had 6+1 rows of delay channels, 6 rows for MNase side, 1 row for MBA side, aliquoting right after droplet generation, "single layer" or thin, 40-µm delay for MBA side, "dual layer" or thick, 160-µm delay for MNase side.

Quantitative PCR as Quality control of Droplet Microfluidic Produced Samples

Digestion efficiency was validated among literature as a criterion for assessing digested chromatin samples' efficacy for MNase-seq[28, 29]; however, it was not the case for accessibility assay, especially for a new protocol as MBA-seq. A more biologically relevant benchmark should

be evaluated to confirm whether accessible regions were enriched in MBA DNA compared to MNase DNA (the latter should contain little to no nucleosome-free, accessible regions). Therefore, qPCR on known accessible and inaccessible regions for fold change calculation was selected to serve as the biologically relevant quality control method. The two known loci used as positive targets that were supposed to be enriched in MBA DNA were one human *CD47 enhancer* and human *DZIP3*. CD47 is a cell surface glycoprotein that inhibits phagocytosis of cells expressing it via binding to its receptor on macrophages and other immune cells. Therefore, CD47 is overexpressed on cancer cells including K-562 cells to survive the guard from immune system and its enhancers are accessible to realize the upregulation.[30] *DZIP3* (*DAZ interacting zinc finger protein 3*) is broadly expressed among multiple tissues and the coded protein is involved in ligase activity, therefore it is considered housekeeping gene and should always be assessible. Meanwhile, an inaccessible region, *C19 intergenic*, was used as a reference region to normalize the fold of positive targets to. This region was discovered to be between two genes, containing a nucleosome structure that has no histone modifications, hence should only exist with significant amount in MNase DNA but little should be found in MBA DNA. Based on empirical observations from collaborators at Mayo Clinic, Rochester, fold change of positive locus normalized against reference locus in MBA DNA should be ~ 5 times of that in MNase DNA, i.e., the result of equation 4.1 ($\Delta\Delta C_t$) should be around or more than $\log_2 5$ or 2.32.

With the expected results bared in mind, qPCR was carried out on amplified and size selected MBA DNA and paired MNase DNA. Results in **Figure 4-5** revealed that except for the samples from the 8RABD device and one trial of device 6RABD, all samples showed fold change more than 5 on both positive targets, *DZIP3* showing higher enrichment compared to *CD47 enhancer*. The two designs that showed expected digestion patterns on both MBA and MNase

sides, 6RABD and 6RASD, delivered samples that had higher enrichment of *DZIP3* compared to other trials and significant enrichment of *CD47 enhancer* across most trials. Considering the bigger absolute values of the fold change and the higher consistency among all trials, the device 6RASD was selected as the optimal device for biological replicate collection from K-562 cells and buffy coat from healthy humans for the subsequent experiments.

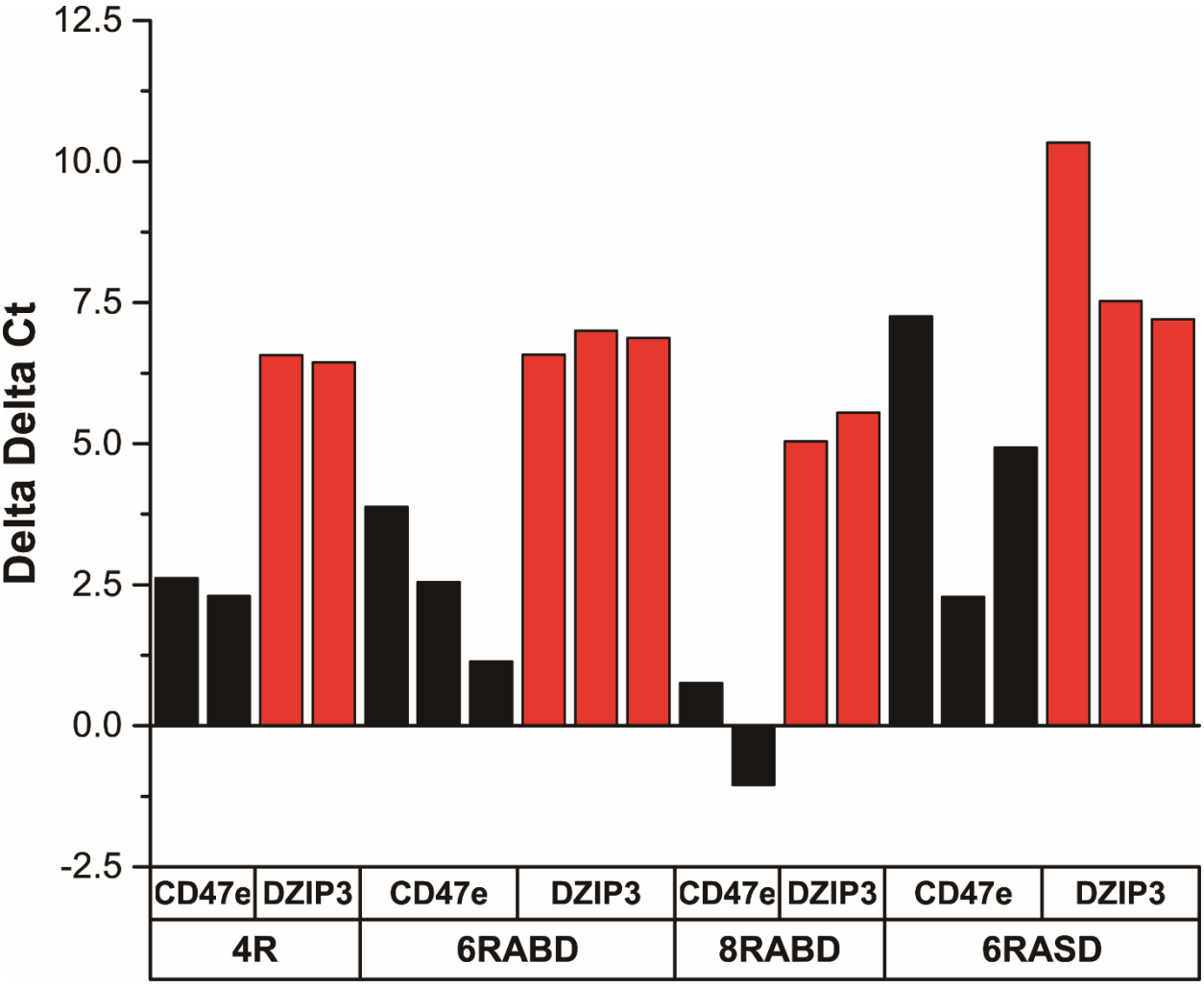

Figure 4-5 Delta Delta Ct ($\Delta\Delta C_t$, or $\log_2$(fold change)) of samples produced from different devices. "CD47e" stood for the targeted region of *CD47 enhancer*, and "DZIP3" stood for the targeted region of *DZIP3*.

K-562 Cells and Buffy Coat Processing for Sequencing

With an optimal design selected, samples from K-562 cells and buffy coat were processed under developed conditions and submitted to the collaborators for next-generation sequencing to obtain a final validation on the efficacy of the developed MBA-seq protocol and the applicability of the droplet microfluidic platform for simultaneous sample collection for nucleosome-free region identification and nucleosome positioning analysis. Figure 4-6 to 9 confirmed the consistent performance of the droplet microfluidic device on biological triplicates for both sample types. Enough DNA were collected on both sides among all replicates to ensure successful library construction (**Figure 4-6** and **8**). More importantly, limited digestion occurred on MBA DNA while the main product for MNase DNA were of mononucleosome size, indicating the high possibility of correct enrichment of nucleosome-free regions and mononucleosomal DNA, respectively (**Figure 4-7** and **9**, also see **Figure B-2 to 5** in **APPENDIX B** for representative K-562 cells and buffy coat digestion patterns characterized with Bioanalyzer). The success of processing buffy coat samples was particularly impactful as buffy coat is a great representative of clinical samples from whole blood, a common medical sample from patients. Buffy coat being successfully processed by the developed droplet microfluidic platform enhanced the potential of this automated platform to be involved in clinical studies and contribute to disease-related information discovery.

One thing to be noticed was that, I started K-562 cell sample collection with 60 µL of cell suspension at a density of 4 million per mL for the first two replicates and increased the volume to 100 µL for the next four replicates to obtain enough MBA DNA for subsequent processing in order to ensure multiple trials can be operated on the same sample if necessary. Similar input increase also existed in buffy coat collection. As buffy coat content was not necessarily the same from different subjects, the final MBA and MNase DNA amount were uneasy to anticipate,

therefore I started with 100 µL of buffy coat suspension at a density of 4 million per mL and increased the cell density to 10 million per mL with the other three replicates to increase the final DNA amount. The DNA amount of these samples with smaller input sizes were corrected proportionally to match with the cell input of other replicates in the corresponding sample pools (K-562 and buffy coat, respectively), and the results were surprisingly consistent across all replicates: none of the corrected values were outliers based on the box plots' interquartile ranges (outliers are the ones bigger than 1.5 × interquartile range above the third quartile or the ones smaller than 1.5 × interquartile range below the first quartile), nor were they away from the means by more than 2 times of the standard deviations. Hence, this observation shed light again on the promising feature of droplet microfluidic platform's sample size flexibility: as long as the cell suspension density did not surpass the upper limit where every generated droplet would contain one cell averagely, the sample size of a droplet microfluidic device could be enlarged by increasing the volume so the platform was run for a longer time, or increasing the suspended cell density so more droplets would contain a cell, without affecting the processing performance on chip as long as the droplets were processed under the same condition on the same (designed) device. This feature can be very practical and user-friendly when such a droplet microfluidic platform is commercialized and applied to treat clinical samples where the sample sizes are more likely to vary among different sources, and it still allows the expected, consistent performance from an automated platform processing samples of significance.

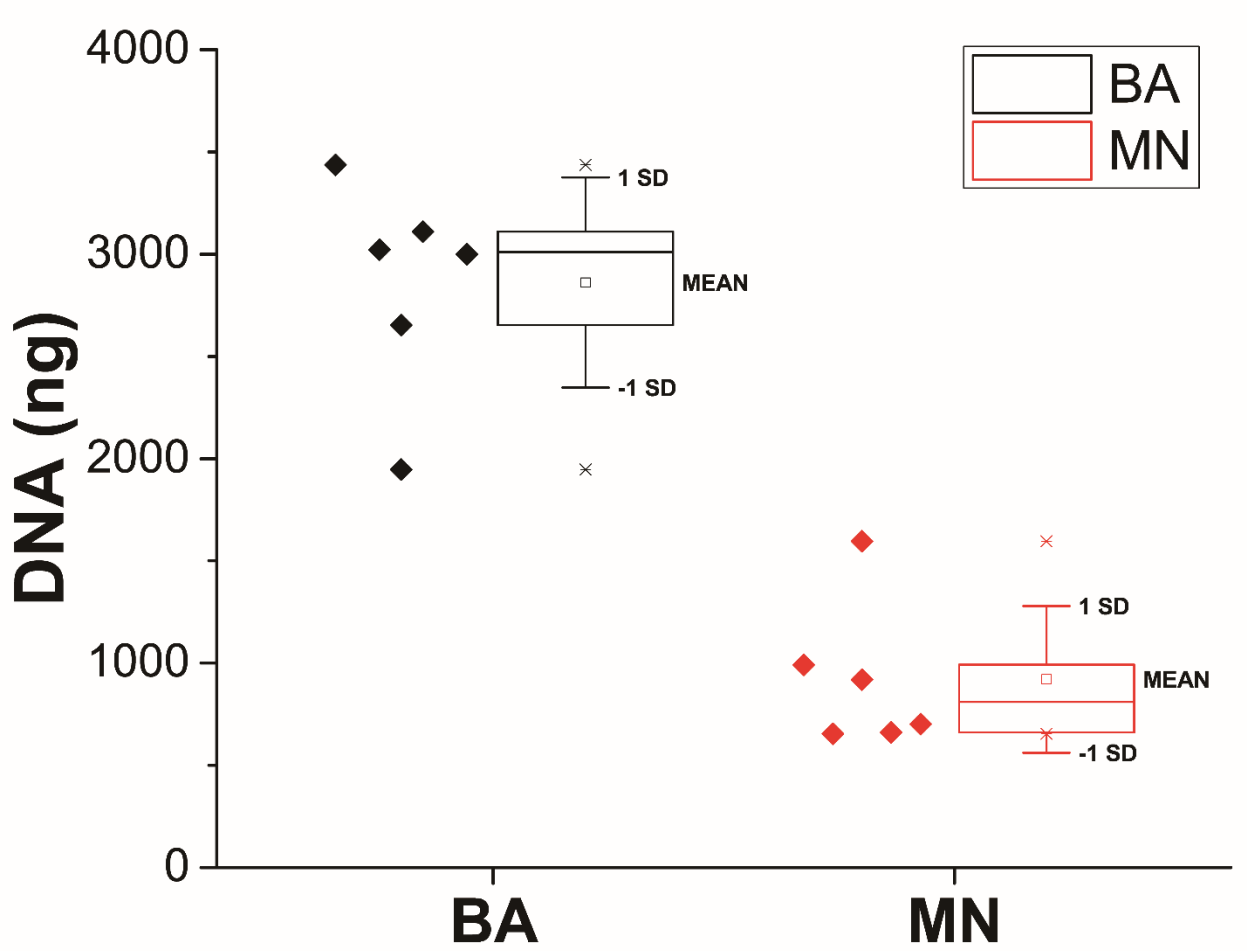

Figure 4-6 K-562 cell samples' DNA content collected from both MBA ("BA") and MNase ("MN") sides. Six replicates were processed. The upper end of the box represented the third quartile of the data group, and the lower end representing the first quartile. The whiskers represented one time of the standard deviation away from the mean from both directions, with mean represented by the square inside the box. Maximum and minimum of the data group were marked by asterisks above and below each box.

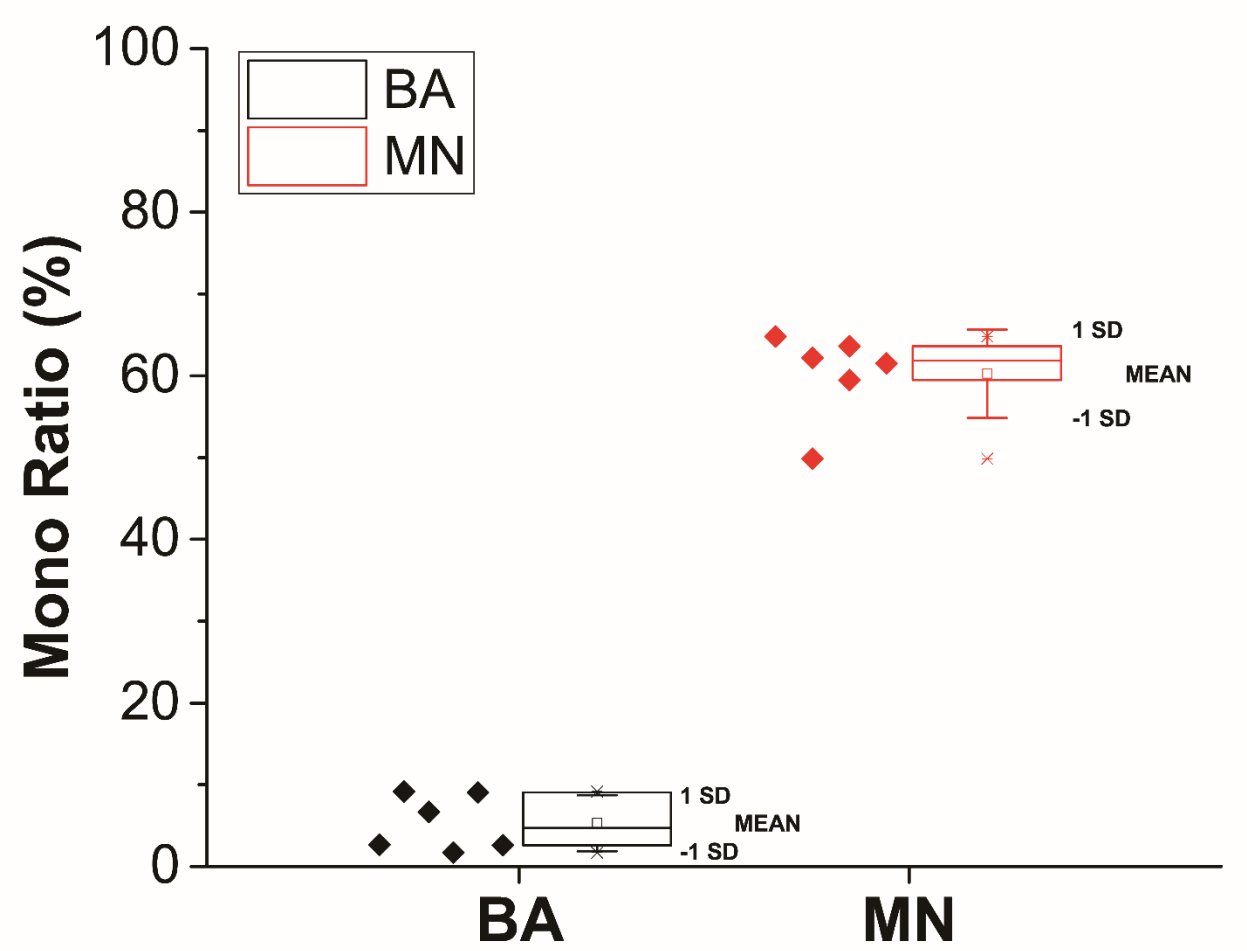

Figure 4-7 K-562 cell samples' mononucleosome (or area below the peak curve about < 250 bp) yields from both MBA ("BA") and MNase ("MN") sides. Six replicates were processed. The upper end of the box represented the third quartile of the data group, and the lower end representing the first quartile. The whiskers represented one time of the standard deviation away from the mean from both directions, with mean represented by the square inside the box. Maximum and minimum of the data group were marked by asterisks above and below each box.

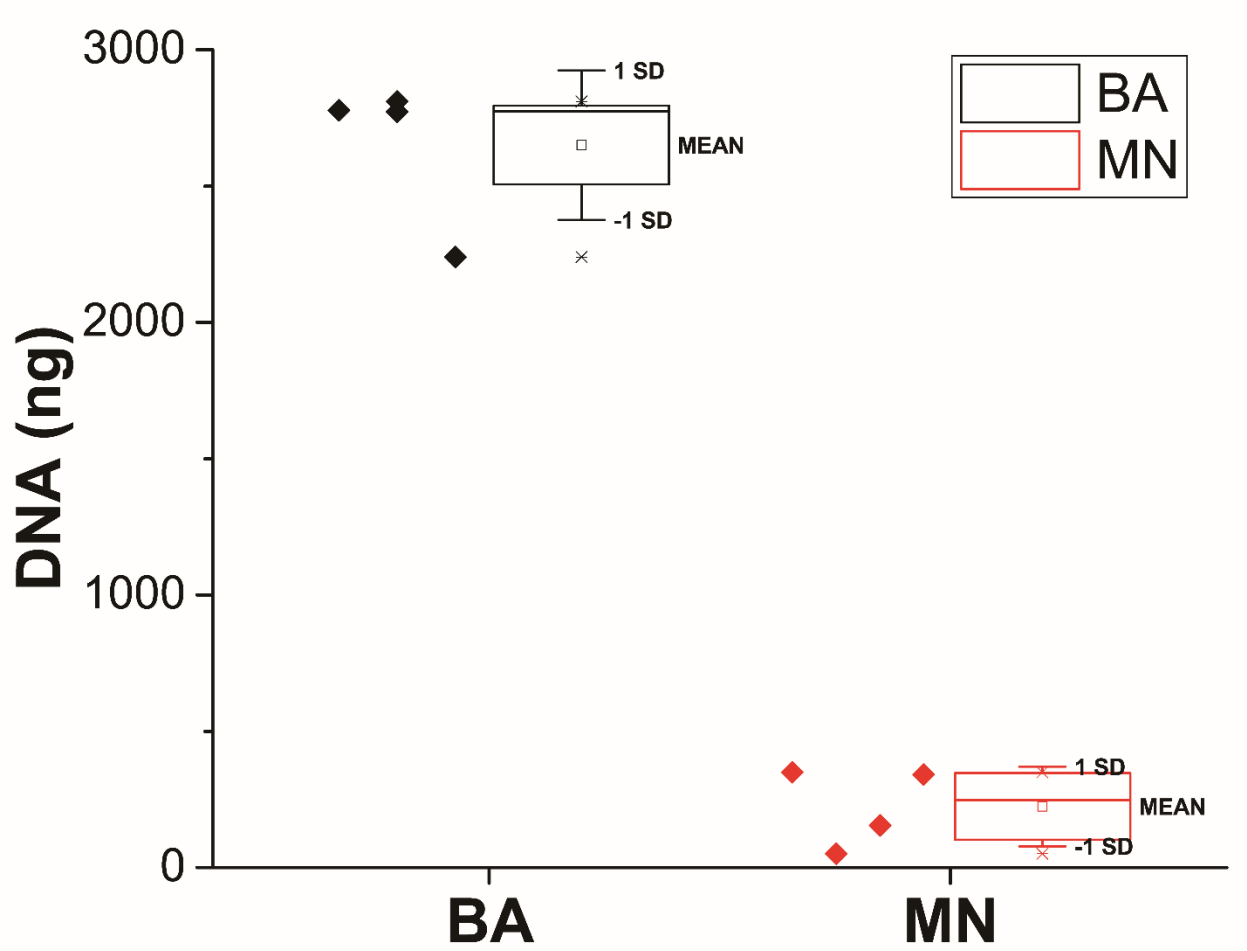

Figure 4-8 Buffy coat samples' DNA content collected from both MBA ("BA") and MNase ("MN") sides. Four replicates were processed. The upper end of the box represented the third quartile of the data group, and the lower end representing the first quartile. The whiskers represented one time of the standard deviation away from the mean from both directions, with mean represented by the square inside the box. Maximum and minimum of the data group were marked by asterisks above and below each box.

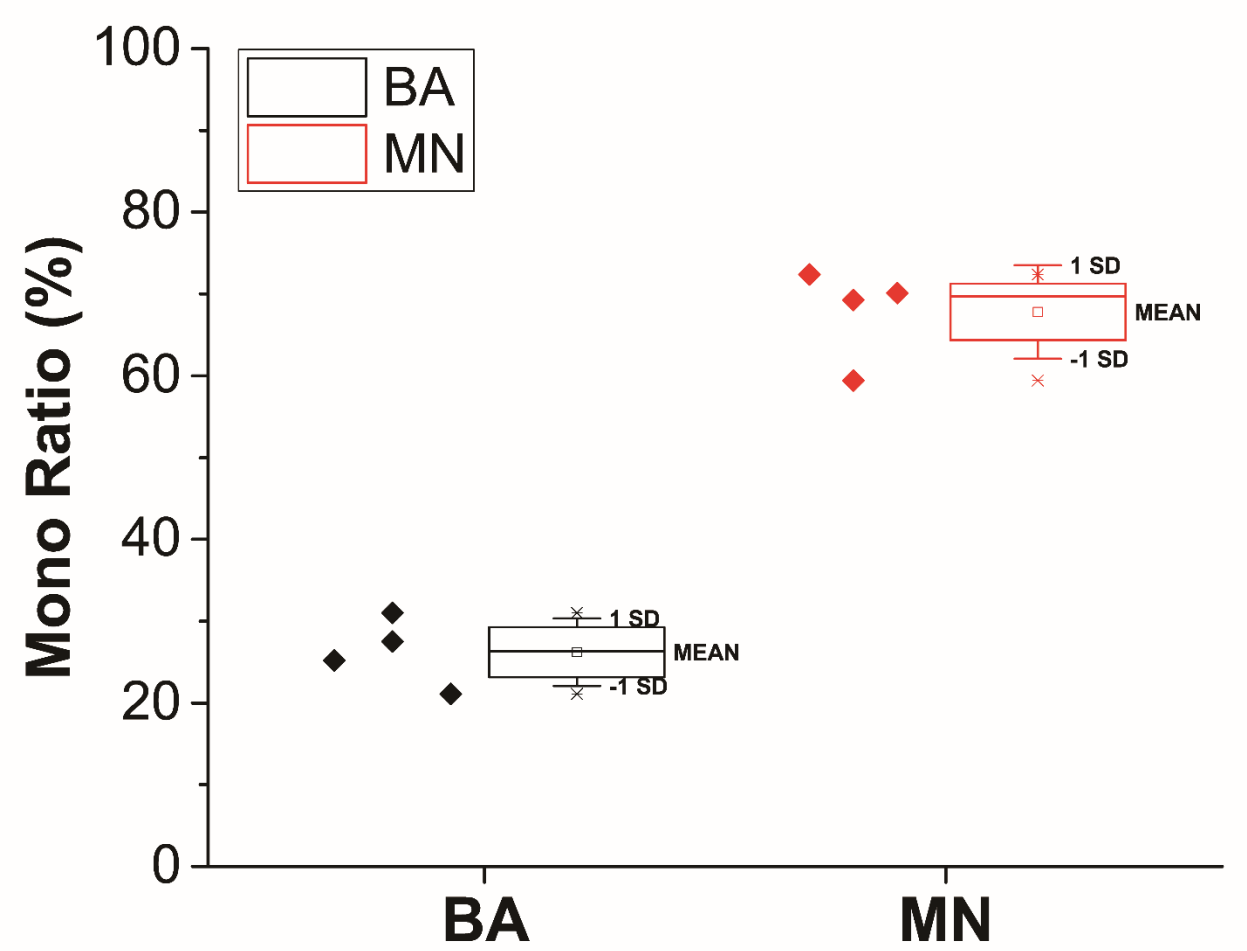

Figure 4-9 Buffy coat samples' mononucleosome (or area below the peak curve about < 250 bp) yields from both MBA ("BA") and MNase ("MN") sides. Four replicates were processed. The upper end of the box represented the third quartile of the data group, and the lower end representing the first quartile. The whiskers represented one time of the standard deviation away from the mean from both directions, with mean represented by the square inside the box. Maximum and minimum of the data group were marked by asterisks above and below each box.

Conclusion

In this chapter, a droplet microfluidic device automating sample preparation for both a newly developed, accessible DNA region assay and a well-established, MNase-seq assay was described. Digestion patterns and qPCR quality controls on known accessible, nucleosome-free regions confirmed the efficacy of using droplet microfluidic devices to deliver qualified fragment chromatin for the investigation of nucleosome-free, regulatory region and nucleosome positioning

simultaneously from the same sample. Samples of a model cell line and samples of clinical significance were prepared with the optimal designed device for the ultimate validation by next-generation sequencing. The development of this device and the future possibility of running multiple of them in parallel for high throughput sample processing provide a splendid potential for the discovery of crucial regulatory regions and related functions in medical-relevant samples, the identification of nucleosome-free regulatory regions' interplay with nucleosomes, and the crosscheck of data reliability and overlay of multiple layers of epigenetic profiles from the two important epigenomic assays (MBA-seq and MNase-seq) that provide complimentary information with each other, and will promote the development of automated epigenetic methodology for the insight into the complex epigenetic regulating network of gene expression.

References

1. W. K. M. Lai and B. F. Pugh, *Nature reviews. Molecular cell biology*, 2017, **18**, 548-562.
2. S. H. Stricker, A. Koferle and S. Beck, *Nature reviews. Genetics*, 2017, **18**, 51-66.
3. C. T. Ong and V. G. Corces, *EMBO reports*, 2012, **13**, 423-430.
4. N. H. Nguyen and J. J. Cheong, *Biochemical and biophysical research communications*, 2018, **499**, 1039-1043.
5. C. Jiang and B. F. Pugh, *Nature reviews. Genetics*, 2009, **10**, 161-172.
6. K. Maehara and Y. Ohkawa, *Scientific reports*, 2016, **6**, 19620.
7. X. Chai, S. Nagarajan, K. Kim, K. Lee and J. K. Choi, *PLoS genetics*, 2013, **9**, e1003778.
8. O. Bell, V. K. Tiwari, N. H. Thoma and D. Schubeler, *Nature reviews. Genetics*, 2011, **12**, 554-564.
9. J. Mieczkowski, A. Cook, S. K. Bowman, B. Mueller, B. H. Alver, S. Kundu, A. M. Deaton, J. A. Urban, E. Larschan, P. J. Park, R. E. Kingston and M. Y. Tolstorukov, *Nature communications*, 2016, **7**, 11485.
10. V. I. Risca, S. K. Denny, A. F. Straight and W. J. Greenleaf, *Nature*, 2016, DOI: 10.1038/nature20781.
11. M. de Dieuleveult, K. Yen, I. Hmitou, A. Depaux, F. Boussouar, D. Bou Dargham, S. Jounier, H. Humbertclaude, F. Ribierre, C. Baulard, N. P. Farrell, B. Park, C. Keime, L. Carriere, S. Berlivet, M. Gut, I. Gut, M. Werner, J. F. Deleuze, R. Olaso, J. C. Aude, S. Chantalat, B. F. Pugh and M. Gerard, *Nature*, 2016, **530**, 113-116.
12. M. R. Corces, J. D. Buenrostro, B. Wu, P. G. Greenside, S. M. Chan, J. L. Koenig, M. P. Snyder, J. K. Pritchard, A. Kundaje, W. J. Greenleaf, R. Majeti and H. Y. Chang, *Nature genetics*, 2016, **48**, 1193-1203.

13. L. Song and G. E. Crawford, *Cold Spring Harbor protocols*, 2010, **2010**, pdb prot5384.
14. A. P. Boyle, S. Davis, H. P. Shulha, P. Meltzer, E. H. Margulies, Z. Weng, T. S. Furey and G. E. Crawford, *Cell*, 2008, **132**, 311-322.
15. H. H. He, C. A. Meyer, S. S. Hu, M. W. Chen, C. Zang, Y. Liu, P. K. Rao, T. Fei, H. Xu, H. Long, X. S. Liu and M. Brown, *Nature methods*, 2014, **11**, 73-78.
16. J. Cooper, Y. Ding, J. Song and K. Zhao, *Nature protocols*, 2017, **12**, 2342-2354.
17. J. D. Buenrostro, P. G. Giresi, L. C. Zaba, H. Y. Chang and W. J. Greenleaf, *Nature methods*, 2013, **10**, 1213-1218.
18. J. D. Buenrostro, B. Wu, H. Y. Chang and W. J. Greenleaf, *Current protocols in molecular biology*, 2015, **109**, 21 29 21-29.
19. J. Wu, B. Huang, H. Chen, Q. Yin, Y. Liu, Y. Xiang, B. Zhang, B. Liu, Q. Wang, W. Xia, W. Li, Y. Li, J. Ma, X. Peng, H. Zheng, J. Ming, W. Zhang, J. Zhang, G. Tian, F. Xu, Z. Chang, J. Na, X. Yang and W. Xie, *Nature*, 2016, **534**, 652-657.
20. J. Xu, A. C. Carter, A. V. Gendrel, M. Attia, J. Loftus, W. J. Greenleaf, R. Tibshirani, E. Heard and H. Y. Chang, *Nature genetics*, 2017, DOI: 10.1038/ng.3769.
21. P. W. Koh, R. Sinha, A. A. Barkal, R. M. Morganti, A. Chen, I. L. Weissman, L. T. Ang, A. Kundaje and K. M. Loh, *Scientific data*, 2016, **3**, 160109.
22. P. G. Giresi, J. Kim, R. M. McDaniell, V. R. Iyer and J. D. Lieb, *Genome research*, 2007, **17**, 877-885.
23. P. G. Giresi and J. D. Lieb, *Methods*, 2009, **48**, 233-239.
24. L. Song, Z. Zhang, L. L. Grasfeder, A. P. Boyle, P. G. Giresi, B. K. Lee, N. C. Sheffield, S. Graf, M. Huss, D. Keefe, Z. Liu, D. London, R. M. McDaniell, Y. Shibata, K. A. Showers, J. M. Simon, T. Vales, T. Wang, D. Winter, N. D. Clarke, E. Birney, V. R. Iyer, G. E. Crawford, J. D. Lieb and T. S. Furey, *Genome research*, 2011, **21**, 1757-1767.
25. M. Tsompana and M. J. Buck, *Epigenetics & chromatin*, 2014, **7**, 33.
26. L. Mazutis, J. Gilbert, W. L. Ung, D. A. Weitz, A. D. Griffiths and J. A. Heyman, *Nature protocols*, 2013, **8**, 870-891.
27. T. H. Ho, R. N. Nateras, H. Yan, J. G. Park, S. Jensen, C. Borges, J. H. Lee, M. D. Champion, R. Tibes, A. H. Bryce, E. M. Carballido, M. A. Todd, R. W. Joseph, W. W. Wong, A. S. Parker, M. L. Stanton and E. P. Castle, *PloS one*, 2015, **10**, e0132831.
28. J. M. Rizzo, J. E. Bard and M. J. Buck, *BMC molecular biology*, 2012, **13**, 15.
29. J. Rodriguez, J. N. McKnight and T. Tsukiyama, *Current protocols in molecular biology*, 2014, **108**, 21 28 21-16.
30. P. A. Betancur, B. J. Abraham, Y. Y. Yiu, S. B. Willingham, F. Khameneh, M. Zarnegar, A. H. Kuo, K. McKenna, Y. Kojima, N. J. Leeper, P. Ho, P. Gip, T. Swigut, R. I. Sherwood, M. F. Clarke, G. Somlo, R. A. Young and I. L. Weissman, *Nature communications*, 2017, **8**, 14802.

Chapter 5 Using Droplet Microfluidics for Automated Chromatin Immunoprecipitation

Abstract

The gold standard to probe histone modification marks as well as other DNA-binding proteins is a technique called chromatin immunoprecipitation (ChIP), commonly followed by next-generation sequencing to complete the protocol of ChIP-seq. While powerful, traditional ChIP protocols require a large cellular input (10^6-10^7 cells), which limits their utility to study biopsies, rare cells such as cancer and stem cells, and to assess tumor heterogeneity. ChIP is also laborious, time-consuming, and highly user-dependent. To address these disadvantages, we proposed a droplet microfluidic strategy to automate from cell lysis to immunoprecipitation of interested marks. The resulting product was clean chromatin fragments ready for DNA extraction and analysis via qPCR or next generation sequencing. This proposed strategy will provide unprecedented opportunity to study epigenetic profiles of low-input to single-cell level samples, allowing both the assessment of cell heterogeneity within complex clinical samples and the application of cost-effective epigenetic testing to very small samples in individualized medicine settings directly at the point of care. This chapter first described the characterization and validation of ChIP efficiency using in-house conditions mimicking the procedure occurring on the microfluidic device, then showcased the results of on-chip ChIP trials. Future directions for optimal performance were pointed out at the end of the chapter.

Introduction

Histones are significant components of nucleosomes and chromatin, assisting in packaging 6 billion bp of human DNA into a micron-sized nucleus structure.[1] H2A, H2B, H3, and H4 are the four core histones that form the octamer wrapped by DNA sequences. They all have flexible N-terminal tails, but also C-terminal tails of H2A, protruding outside the nucleosomes.[2] Most regions on these tails are potential targets of a variety of chemical modifications including acetylation, methylation, phosphorylation, and others, which are one category of the most investigated epigenomic regulators.[3] They are recognized by specific "reader" protein complexes to trigger subsequent processes and are discovered to be associated with the regulation of fundamental processes such as initiation of transcription, activation/repression of gene expression and regulatory regions, and remodeling of chromatin conformation, thus are crucial in cell development, aging, and carcinogenesis.[4-6] For example, trimethylation of H3K4 (H3K4me3) has been commonly found at transcription start sites (TSSs) of actively transcribed genes and is discovered to improve transcription machinery recruitment[7], respond to induce gene expression upon environmental cues[8], and maintain transcription consistency[9]. Mono- and di-methylation on H3K4 (H3K4me1 and H3K4me2, respectively) are marks for enhancers to help establish *de novo* enhancers and poise them for activation.[10, 11] Acetylation of H3K27 (H3K27ac), after established by histone acetyltransferase, is associated with enhancer activation to promote gene expression, and are often found overlapped with H3K4me1.[12, 13] On the other hand, H3K27me3 and H3K9me2 are associated with repressed gene expression.[14-16] Additionally, it has been recognized that histone marks usually work in combination with each other and other epigenetic marks to form a regulating network.[14, 17] Their deposition and removal from chromatin is part of the complicated interplay between epigenome and genome, necessitating the persistent efforts trying to comprehend the

locations and functions of histone modifications in different cell statuses such as healthy vs. diseased and among different cellular development stages to fully clarify their significance in epigenetics.[18, 19]

Currently, the standard method to localize histone modifications is the protocol called chromatin immunoprecipitation (ChIP), frequently followed by next-generation sequencing (ChIP-seq).[1] ChIP-seq generally involves steps of shearing chromatin into mono- and di-nucleosome sizes (about 150-300 bp) by sonication and/or micrococcal nuclease, immunocapturing the modification of interest with specific antibodies, and purifying captured DNA for library preparation and sequencing.[20-22] ChIP is an affinity-based assay, therefore it can selectively target other DNA-binding proteins besides histones simply by using appropriate antibodies and chemically fixing chromatin with interacting proteins to preserve transient interaction prior to chromatin fragmentation.[23] Thus, ChIP is also known as the gold standard to investigate DNA-protein interactions from a broad aspect to localize those proteins involved in epigenetic regulation (e.g., those recruited by histone modifications).[24]

ChIP-seq is indeed a powerful tool in elucidating epigenetic regulating network. With ChIP-seq and its derivative methods, identification and validation of modifiers and regulators related to aging, cellular development, disease/cancer progression, and other crucial biological processes have been realized, along with the disclosure of underlying epigenetic regulating mechanisms.[25-30] However, it has several limitations that people want to improve upon. One major challenge is the large input sample size requirement. Traditionally, ChIP-seq requires more than 10^6 cells as starting materials for one mark to ensure effective enrichment of DNA related to protein of interest, which is unfeasible for clinical biopsies containing fewer than 10^3 cells each or samples of rare cell populations.[22, 31] A second challenge is the lack of sensitivity to sample

heterogeneity due to the aggregation of DNA from all cells present in the sample of interest. Cells are usually lysed, digested, and analyzed in bulk, which buries possible signals from small numbers of abnormal or novel features. Third, robustness and reproducibility of ChIP-seq need to be enhanced. ChIP-seq usually relies largely on an operator's skill to complete a large number of manual processing steps, thus the results can vary to a large extent simply depending on qualification of the operator. Fourth, ChIP-seq experiments are inherently low throughput, where they target only one protein modification per sample volume. When combined, these factors complicate the use of conventional ChIP-seq in routine clinical practices, such as those studying large patient cohorts including monitoring both healthy and abnormal samples across human life span.[32] Efforts have been made to improve ChIP-seq and typical examples are those using continuous microfluidics as the latter has been widely known to offer the advantages of automation and reduced sample sizes and reagent consumption. The Quake group has been one of the pioneers in enabling microfluidic low-input cells and single-cell epigenomic sequencing and profiling. They have developed automated microfluidic device-based ChIP protocol (AutoChIP) that allows 2000-cell equivalent DNA to be enriched with similar quality of ChIP DNA compared to conventional methods.[33, 34] Multi-layered and valve-assisted ring structures were designed to control sample loading, bead washing, and DNA eluting. By integrating multiple AutoChIP structures, they have also improved the throughput of ChIP experiments with 10,000-cell equivalent chromatin as starting material per ring.[35] The Lu lab has also been investigating epigenomic assays for fewer cells by exploiting microfluidic chamber structures. By using valves to control cell lysate loading and capture DNA elution, they have shown that fragmented chromatin from 100 cells can be treated microfluidically with high enough quality for subsequent DNA analysis.[36] Further integration of sonication and immunoprecipitation enabled on-chip shearing of chromatin in

addition to DNA capture.[37] Their most recent publication also showed that by fabricating multiple microfluidic beds ChIP targeting two marks could be carried parallelly.[38] These improvement were all enabling more capable ChIP on smaller sample sizes; however, they all rely on continuous microfluidics which the device size becomes the limiting factor of the sample size that is compatible with the device: only a fixed amount of fragmented chromatin that fits the size of the microfluidic chamber, and fits with the amount of beads functionalized with antibodies that are (pre-)loaded into the chamber can be processed, too big sample size leading to sample volume pushed out of the chamber thus sample loss, too small size leading to a mismatch with the number of antibody-functionalized beads thus deviating from the optimized ChIP efficiency. Additionally, not all microfluidic ChIP starts with cells directly, some using pre-fragmented chromatin as starting materials. Hence, space for enhancement still exists for ChIP automation and minimization, and droplet microfluidics is a potential tool for such optimized ChIP.

Droplet microfluidics is becoming a valuable tool for various applications such as single-cell analysis, complex biological and chemical assays, diagnostics, DNA sequencing, and drug screening.[39-42] The basic principle to generate droplets includes segmenting aqueous streams with a continuous, immiscible, and inert oil flow, and each droplet functions as an independent microreactor surrounded by oil and not touching the inner side of the channel, indicating little nonspecific adsorption of samples to the device or contamination with each other.[43, 44] The volume of the droplets is tunable by changing flow rate values and ratios between continuous and dispersed phases, and typically ranges from femtoliters to nanoliters thus suitable for single cell analysis.[45-58] Reliable and precise manipulation of individual droplet has emerged recently covering droplet coalescence, splitting, injection, and content mixing.[59-69] Additionally, previous studies have shown that mixing occurs very rapidly within the droplets due to convective flow profile inside

them, allowing for more efficient chemical and biological reactions and thus shorter operation time for experiments.[43, 70-72] This is highly advantageous over conventional continuous-flow microfluidics, which offers low Reynolds numbers and laminar flows, resulting in rather slower mixing predominated by molecular diffusion, as it means droplet microfluidic devices need little to no external mechanisms to help mixing, leading to the simplification of device design and operation.[43] Furthermore, as droplets can be continuously generated and cell distribution in droplets follow Poisson distribution, a flexible sample size can be processed on a droplet microfluidic device without an affected processing performance by keeping it operating until all the volume is segmented into droplets as long as the cell suspension density does not surpass the theoretical upper limit, that is one cell per generated droplet. This flexible sample size feature is practically useful for clinical sample ChIP, as knowing the specific cell numbers per sample may be impossible in real life.

With all the advantages discussed, a droplet microfluidic strategy to automate and improve ChIP was described in detail in this chapter. Using a slightly modified cellular chromatin MNase processing device adapted from Chapter 3, cells were lysed and digested simultaneously inside droplets and antibody-functionalized beads suspended in quenching buffer were injected into these droplets at the end of the device, allowing ChIP to happen. The collected droplets were loaded to a second device to wash the chromatin-antibody-bead conjugates to reduce non-specific binding before eluting and extracting the bound DNA. The majority of the chapter discussed the validation of ChIP under the droplet microfluidic conditions, with preliminary data of droplet microfluidic ChIP briefly commented. Quantitative PCR on gene loci associated with the targeted histone modification marks was executed to evaluate the ChIP enriching efficiency of the chromatin

regions containing tested marks. Directions for further optimization to improve ChIP in flowing droplet microfluidic format were listed and explained in the end.

Materials and Methods

Buffer recipes

Lysis buffer (pH 7.9) contained 10 mM HEPES, 1.5 mM magnesium chloride ($MgCl_2$), 10 mM potassium chloride (KCl), and 0.5% (w/v) IGEPAL-CA630 (Sigma-Aldrich, St. Louis, MO). Digestion buffer (pH 7.5) contained 20 mM Trizma hydrochloride (Tris-HCl), 15 mM sodium chloride (NaCl), 60 mM KCl, 5 mM calcium chloride ($CaCl_2$), 0.15 mM spermine, and 0.5 mM spermidine. 2× STOP/ChIP buffer or quenching buffer (pH 8) contained 100 mM Tris-HCl, 20 mM ethylenediaminetetraacetic acid (EDTA), 200 mM NaCl, 2% Triton X-100, and 0.2% sodium dodecyl sulfate (SDS). Low salt immune complex wash buffer (pH 8.1) contained 20 mM Tris-HCl, 2 mM EDTA, 150 mM NaCl, 0.1% SDS, and 1% Triton X-100. High salt immune complex wash buffer (pH 8.1) contained 20 mM Tris-HCl, 2 mM EDTA, 500 mM NaCl, 0.1% SDS, and 1% Triton X-100. LiCl immune complex wash buffer (pH 8.1) contained 10 mM Tris-HCl. 1 mM EDTA, 250 mM LiCl, 1% IGEPLA-CA630, 1% sodium deoxycholate. TE buffer (pH 8.1) contained 10 mM Tris-HCl and 1 mM EDTA. 2×elution buffer (pH 8.0, recipe from Mayo Clinic, Rochester) contained 20 mM Tris-HCl, 20 mM EDTA, 300 mM NaCl, and 2% SDS.

Cell culture and cell suspension preparation

HeLa cells were cultured at 37°C with 5% CO_2 in the DMEM medium (Gibco DMEM high glucose, Catalog No. 11965092, Fisher Scientific) supplemented with 10% v/v fetal bovine serum.

After aspirating existing media and washing with cold PBS, the cells were treated with 3-mL, 0.5% trypsin-EDTA for 5 minutes at 37°C and added 7-mL complete growth media to quench trypsin's activity. Then the cell suspension was transferred to a 50-mL conical tube and pelleted at 800×g for 5 min at 4°C, washed with cold PBS buffer, and resuspended into the density-adjusted PBS buffer to a defined concentration as described below. After resuspension, cells were stored on ice for use. All digestion and ChIP experiments of HeLa cells (both on chip and in tube) were carried out with 30 µL of cell suspension at a density of 4 million cells per mL, unless otherwise stated.

Immunoprecipitation bead functionalization

For each ChIP reaction, 30-µL protein G Dynabead slurry for immunoprecipitation (Invitrogen, Catalog No. 10003D and 10004D, Thermo Fisher Scientific) was washed by 100-µL 1× STOP/ChIP Buffer for a total of three times. Washed Dynabeads was resuspend in 100-µL 1× STOP/ChIP Buffer. A specific volume of antibody stock solution was added to the beads for an expected amount. For example, 10-µL, 0.2 mg/mL anti-H3K4me3 antibody (Anti-Histone H3 (trimethyl K4) antibody-ChIP Grade, Catalog No. ab8580, Abcam) was added to the 100-µL bead suspension for a ChIP reaction with 2-µg antibody. Beads were incubated with antibodies at 4 °C on a rotator for at least 2 hours. After removal of supernatant, the functionalized beads were washed by 100-µL 1× STOP/ChIP Buffer for a total of three times and supernatant aspirated. The beads were ready for ChIP.

Microfluidic device fabrication

Microfluidic devices were generated using standard photolithography and soft lithography methods with details described in Chapter 2 and the brief summary in Chapter 3 and 4. The pattern of the device was designed in AutoCAD (schematic for the device's design was shown in **Figure 5-1**). Briefly, SU8 2025 (MicroChem Corp, Westborough, MA) was first spin coated (PWM 32, Headway Research, Inc) onto the silicon wafer at 2000 rpm for 30 seconds to give a 40-μm thick layer. After soft baking, the resist layer was exposed to a UV lamp (Optical Associates, Incorporated) through the first-layer photomask. After post-exposure baking, the patterned wafer was developed in propylene glycol monomethyl ether acetate (PGMEA) (Sigma-Aldrich, St. Louis, MO) and dried under nitrogen. SU8 2050 was then spin coated onto the patterned wafer at 1150 rpm for 30 seconds to give a 160-μm thick layer. After soft baking, the resist was exposed to UV through the second-layer photomask and then developed as described above. After that, the PDMS base and elastomer curing agent (Momentive RTV 615 kit, R.S. Hughes, Carol Stream, IL) were mixed at a mass ratio of 10:1 and degassed under vacuum before poured over the master. After curing and peeled off from the master, the PDMS replica was punched for tubing inserting holes with blunt-tip biopsy needles (18 gauge for inlets, 20 gauge for the outlet and electrolyte inlets). The device replicas were then sonicated in water for 5 min and dried with nitrogen to remove any PDMS debris. In parallel, a base layer was created on a glass slide by spinning a degassed, 10:1 PDMS base:elastomer mixture onto the glass microscope slide. This base layer was partially cured for 8 min before the microfluidic replica was placed on top with gentle pressure, followed by a final 30 min bake to fully bond the microfluidic device. Alternatively, the PDMS replica and the microscope glass slide could also be put into a plasma chamber (Plasma cleaner PDC-32G, Harrick Plasma, Inc), treated for 30 s at high level of oxygen plasma with 1 torr of air pressure, and

assembled together by attaching the treated side of the PDMS replica to the treated side of the glass slide. A 30-min bake to fully bond the microfluidic device is still necessary.

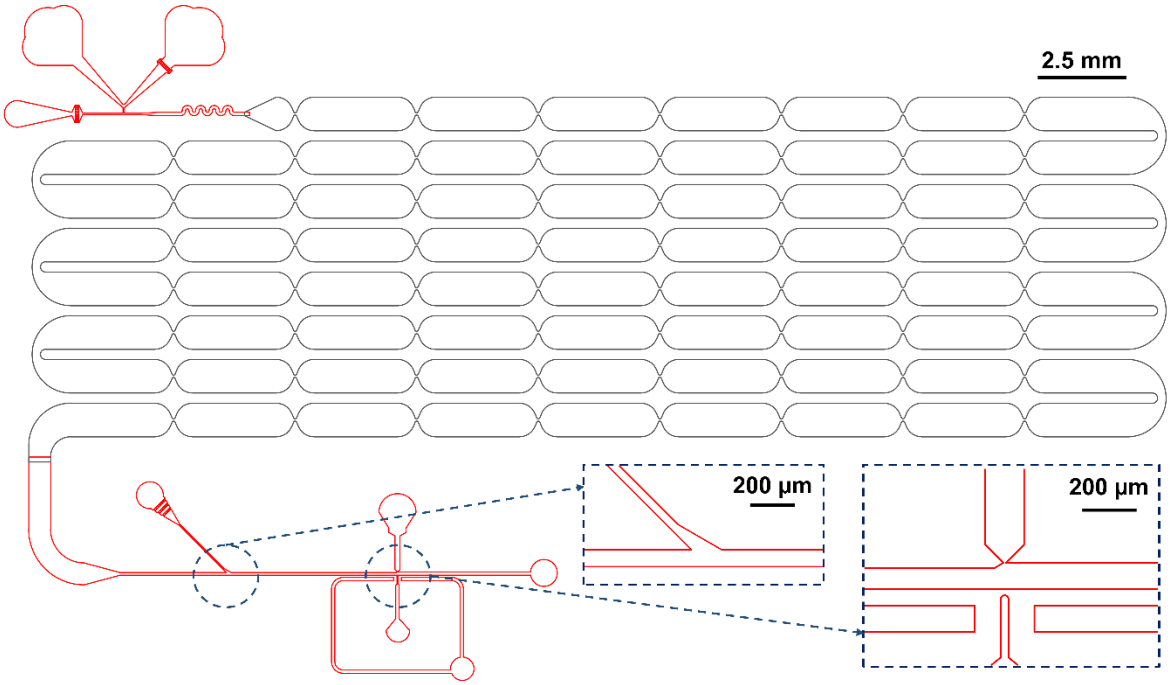

Figure 5-1 Schematic of the photomask design of the droplet microfluidic device for on-chip ChIP. The red parts were the first layer of 40-µm thickness, and the black part the second layer of 160-µm thickness. The two layers were printed on two individual masks. They should be aligned to ensure the connection of the channels (shown as the overlap between the red and black). The zoom-in insertions showed the design for the oil spacer orifice and the modified, widen-up picoinjector, respectively.

Microfluidic device operation

Solutions were delivered to the device through 24-gauge PTFE tubing (Cole-Parmer, Vernon Hills, IL) controlled by syringe pumps (Pump 11 Pico Plus Elite, Harvard Apparatus). The immiscible oil phase was the fluorinated oil Novec 7500 (3M, St. Paul, MN) with 2% (w/w) perfluoropolyether polyethylene glycol block co-polymer surfactant (RAN Biotechnologies, Inc. Beverly, MA). The tubing to deliver cells to the device was treated with 1% (w/v) Pluronic F127 (Sigma-Aldrich, St. Louis, MO)-containing PBS solution overnight and prefilled with PBS buffer.

Two sets of electrolyte channels for the two injection sites on the microfluidic device for MBA and MNase samples respectively were all filled with a 3M sodium chloride solution. For each set, two clamps from electrodes of an in-house built DC-AC converter (12V DC input, 0-360V AC at 36 kHz as output), which was connected to a power supply (DC regulated power supply, TENMA T2-6628), were connected to the syringe needles supplying the electrolyte channels to form a complete circuit for the AC electric field.

Prior to experiments, microfluidic devices were treated with Aquapel (Pittsburgh Glass Works, Pittsburgh, PA) as described previously[53] to give a hydrophobic coating throughout sample processing. Optiprep (Sigma-Aldrich, St. Louis, MO) was added to PBS buffer to obtain a final density of 1.07 g/mL at a 21.9% (v/v) concentration. Cells were resuspended in this density-adjusted PBS buffer to reduce clumping and settling in the tubing. Cell suspensions were drawn into the pre-treated tubing separated from the pre-filled buffer by a plug of air. The combined lysis-digestion buffer was prepared by mixing 30-µL MNase (2000 GU/µL; New England Biolabs, Ipswich, MA), 120-µL digestion buffer, and 300-µL lysis buffer. Antibody-functionalized beads were suspended in the 2×STOP/ChIP buffer at a ratio of 100 µL of buffer per number of beads from 50 µL of original bead slurry mixed with 1 µL of protease inhibitor (Protease inhibitor cocktail set II, Catalog No. 539132, EMD Millipore Corporation). This ratio is the same as that described in bead functionalization section but scaled up the total volume. The syringe for bead containing STOP/ChIP buffer was pre-filled with normal 2×STOP/ChIP buffer so that the 100 µL bead containing quenching buffer was only in the PTFE tubing.

Flow rates of oil (6 µL/min), cell suspension (2 µL/min), lysis-digestion reagent (2 µL/min), 2×STOP/ChIP buffer with antibody-functionalized beads (4 µL/min), and spacing oil (10-14 µL/min) were optimized to ensure that droplets with appropriate size were generated, droplets

travelled across the delay channels for appropriate time to allow thorough MNase digestion, and quenching buffer containing antibody-functionalized beads was injected into every droplet at the injection sites. Droplets were collected via 30-gauge PTFE tubing (Cole-Parmer) into 0.6-mL tubes for the subsequent bead wash. Device operation was visualized using a high-speed camera (Phantom Miro Ex2) mounted onto a stereoscope (Leica M80).

Chromatin immunoprecipitation (ChIP) under multiple conditions

Traditional ChIP was carried out as described below. Per ChIP reaction, 30-µL combined lysis-digestion buffer with MNase was incubated with 30-µL cell suspension for 3.5 minutes and quenched with 60-µL 2×STOP/ChIP buffer. After 5 minutes of incubation, 880-µL 1×STOP/ChIP buffer was added to the fragmented chromatin and mixed thoroughly. 900-µL fragmented chromatin was combined with prepared antibody-functionalized beads with 4.5-µL protease inhibitor and incubated for immunoprecipitation at 4 °C on a rotator overnight. 9-µL fragmented chromatin from the same source was saved as 1% input, mixed with 91-µL 1×elution buffer, and frozen at -20°C until ready for DNA extraction. After overnight incubation, ChIP reaction mix was set on a magnetic rack for targeted chromatin-antibody-bead conjugates washing. After removal of chromatin solution, 1-mL low salt immune complex wash buffer was aliquoted to each ChIP reaction and mixed thoroughly. The reaction was kept on the rotator at 4°C for 5 minutes. Then the wash buffer was removed. Repeat with three other wash buffers to remove non-specific bound chromatin. The washed chromatin-antibody-bead conjugates were resuspended in 100-µL 1×elution buffer. The ChIP sample was ready for DNA extraction and purification.

ChIP in tube mimicking on-chip conditions (also referred to as off-chip) was carried out as described below. Per ChIP reaction, the prepared antibody functionalized beads were suspended in 60-µL 2×STOP/ChIP buffer mixed with 0.6-µL protease inhibitor. 30-µL combined lysis-digestion buffer with MNase was incubated with 30-µL cell suspension for 3.5 minutes and quenched with 60-µL bead containing 2×STOP/ChIP buffer to quench the digestion and start immunoprecipitation in one pot. The reaction was then incubated overnight at 4 °C on a rotator. After overnight incubation, the ChIP reaction mix was washed with the four types of wash buffer and resuspended in 1×elution buffer as described above. One simultaneously lysed and digested chromatin quenched with normal 2×STOP/ChIP buffer (i.e., no beads suspended inside or protease inhibitor) was prepared as 100% input for this format of ChIP.

ChIP on chip was carried out as described in the section of "Microfluidic device operation". Collected droplets were either incubated overnight at 4 °C and manually washed with the four wash buffer solutions (in this case, the collecting tube should be 1.5-mL instead of 0.6-mL as described in the referred section) or washed by a droplet microfluidic device right after collection. After incubation overnight but before manual washing or after device washing, droplets were coalesced by adding 50-µL 1H,1H,2H,2H-perfluoro-1-octanol (Sigma-Aldrich, St. Louis, MO). Aqueous phase with beads and oil phase were separated by centrifugation at 8000 rpm for 1 min at room temperature in a tabletop microcentrifuge (Centrifuge 5418, Eppendorf, Hauppauge, NY). Aqueous supernatant was removed leaving only beads capturing targeted chromatin. If beads were device washed, they would be resuspended in 100-µL 1×elution buffer directly. Otherwise resuspending in 100-µL 1×elution buffer after manual washing. Finally, these beads The ChIP sample was ready for DNA extraction and purification.

ChIP DNA and input DNA extraction and purification

For each 100-µL elution buffer suspended beads or input fragmented chromatin, 2-µL RNase A (10 mg/mL, Sigma-Aldrich, St. Louis, MO) was added to degrade RNA by incubating at 65 °C for 1 hour. 3-µL Proteinase K (10 mg/mL, ThermoFisher Scientific, Grand Island, NY) was added consecutively to degrade proteins by incubating at 65 °C for 2 hours. During these three hours, bound chromatin was also eluting from the beads into the aqueous phase. Aqueous phases were collected to clean DNA LoBind tubes. ChIP DNA and 1% input were extracted and purified by QIAgen MinElute kit, eluted to 15-µL EB buffer, and 100% input DNA was extracted and purified by QIAquick PCR purification kit, eluted to 50-µL EB buffer.

Characterization of input DNA

The concentration of purified DNA was determined using the Qubit dsDNA assay (ThermoFisher Scientific, Grand Island, NY). The size distribution of the input DNA was characterized using a Bioanalyzer (Agilent). All samples were diluted to or below 500 pg/µL consistent with Bioanalyzer sample submission requirements. The percentage of mononucleosomal DNA in the total DNA was obtained from the Bioanalyzer software.

Quality control of ChIP DNA

Three loci (*Brg1_TSS*, *Myt1_TSS*, *SAT-alpha*) associated with different histone modification marks (H3K4me3, H3K27me3, and H3K9me3, respectively) were selected as targets to assess the enrichment efficiency of ChIP. An intergenic locus (C19 intergenic region[73]) was

used as reference evaluating the digested chromatin's background for ChIP. Sequences of primer pairs for these four regions were listed in **Table B-2** in **APPENDIX B**.

Quantitative PCR reaction per well on the 384-well reaction plate (Applied Biosystems Catalog No. 4309848, Fisher Scientific) was set up according to the manufacturer's recommended parameters. Briefly, for each target, 10-µL PowerUp™ SYBR™ Green master mix (Applied Biosystems Catalog No. A25776, Fisher Scientific) was mixed with 2-µL forward primer (5 µM) and 2-µL reverse primer (5 µM). Then 2 to 3 µL of DNA template was added to the mix depending on the DNA concentration to ensure appropriate amount of DNA was used. The volume was made up to 20 µL by adding appropriate volume of nuclease-free water. After setting up the PCR plate, the experiment was run on 7900HT Fast Real-Time PCR system (Applied Biosystems) with the standard parameter setup at the DNA sequencing core facility (University of Michigan). Fold change of targeted loci (posi) between ChIP DNA (IP) and input DNA (Inp) was calculated using equation (**5.1**) to assess the efficiency of ChIP enrichment of chromatin regions containing each histone modification mark.

$$log_2\ fold\ change = c_{t(IP,reference)} - c_{t(IP,posi)} - (c_{t(Inp,reference)} - c_{t(Inp,posi)}) \quad (5.1)$$

Results and Discussion

The development of ChIP protocol usually contains the optimization of the amount ratio among antibody, beads, and fragmented chromatin, and the confirmation of chromatin's fragmentation efficiency.[24, 36, 74] The same parameters were evaluated here as well. As shown in **Figure 5-2**, when keeping the bead amount and fragmented chromatin amount consistent, increasing the amount of antibody would increase the amount of captured ChIP DNA (until all

possible chromatin fragments containing the corresponding mark was pulled down and capture efficiency reaching saturation). Importantly, the fact that significantly smaller amount of DNA was captured when no antibody was functionalized onto the beads ($p < 0.01$ comparing 0-µg antibody groups and 1-µg antibody groups) confirmed the quality of selected beads and protocol (i.e., ChIP in tube mimicking on-chip conditions) as little non-specific binding happened verifying little noise of the reaction. Additionally, the input sample contained 1-µg fragmented DNA, thus the pull-down efficiency of the four tested antibody amounts were 0.22%, 0.56%, 0.66%, and 0.98%, respectively, agreeing with the collaborators' empirical observation at Mayo Clinic, Rochester.

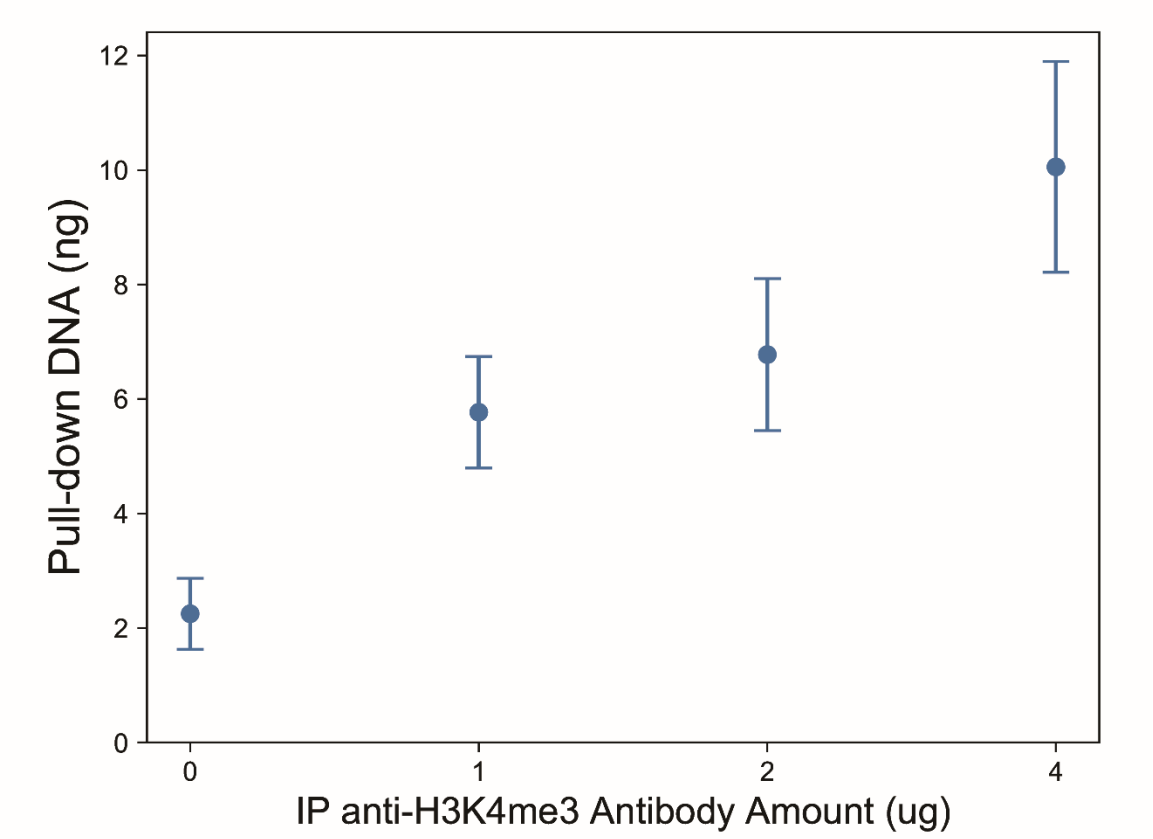

Figure 5-2 ChIP DNA pulled down by various amount of anti-H3K4me3 antibody per reaction. Error bars were standard deviations from three technical replicates under each condition, and each replicate had chromatin digested from individual 30-µL cell suspension aliquot.

After checking the captured ChIP DNA's amount, input samples' digestion patterns should also be checked to ensure expected fragmentation happened. **Figure B-6** confirmed that most chromatin became mononucleosomes after the simultaneous lysis and digestion, with mono ratio at 80.6%. Since the input sample was prepared the same way as those used for ChIP, this result indicated that the fragmented chromatin used for ChIP were mostly mononucleosomes, thus if sequenced, the localization of H3K4me3 mark could be more precisely determined after mapping the reads containing short templates to the reference genome compared to those carried out with bigger fragments.

Besides the captured ChIP DNA and fragmentation of chromatin, qPCR quality control experiments were even more important, and it would confirm the enrichment and the enriching specificity reached by the specific ChIP reaction. Specifically, among the three tested loci (not including the reference locus), only the one within the gene region of Brg1 (*Brg1_TSS*) should be enriched in the ChIP reactions targeting H3K4me3. This is because the encoded product protein Brg1 is a transcription activator, a component of many chromatin remodeling complexes including SWI/SNF, thus should always be activated in HeLa cells and associated with the well-known active mark H3K4me3. On the other hand, *Myt1_TSS* and *SAT-alpha* should not have this mark as the former encodes a fat component dominantly expressed in the nervous system but not cervical originated cells and the latter represents repetitive satellite regions associated with H3K9me3 in centromeres. **Figure 5-3** proved that all three tested antibody amounts enriched significant amount of DNA that once formed into nucleosomes with histone H3 marked by trimethylation on lysine 4. The fold change of *Brg1_TSS* normalized to *C19_intergenic* post compared to prior to ChIP surpassed empirical threshold 5, or $\Delta\Delta C_t$ surpassed 2.32 (i.e., $\log_2$(fold change) in eq 5.1) and were significantly higher than that of the bead control (i.e., 0-µg antibody ChIP) indicating the

concentration of the appropriate chromatin regions containing H3K4me3 by ChIP, while the fold change of *Myt1_TSS* were much lower in all conditions especially in the case of 2-µg antibody per ChIP (the fold change was approximately 1), indicating no significant enrichment of non-specific chromatin regions. Although the fold change of *SAT-alpha* was around the threshold, it was consistent across all conditions including the bead control and thus indicated that perhaps since *SAT-alpha* was repetitive, the multiple copies of that gene contributed to the non-specific pull down into the ChIP DNA and therefore it was not specifically concentrated by the antibody. Together, results in Figure 5-3 validated that the ChIP conditions mimicking on-chip ChIP allowed successful selective enrichment, which strengthened the confidence of the success in droplet microfluidic ChIP development. 2-µg anti-H3K4me3 antibody was selected as the optimal amount for this mark because of its significant enrichment of associated gene locus but lowest non-specific pull down of both non-associated gene loci.

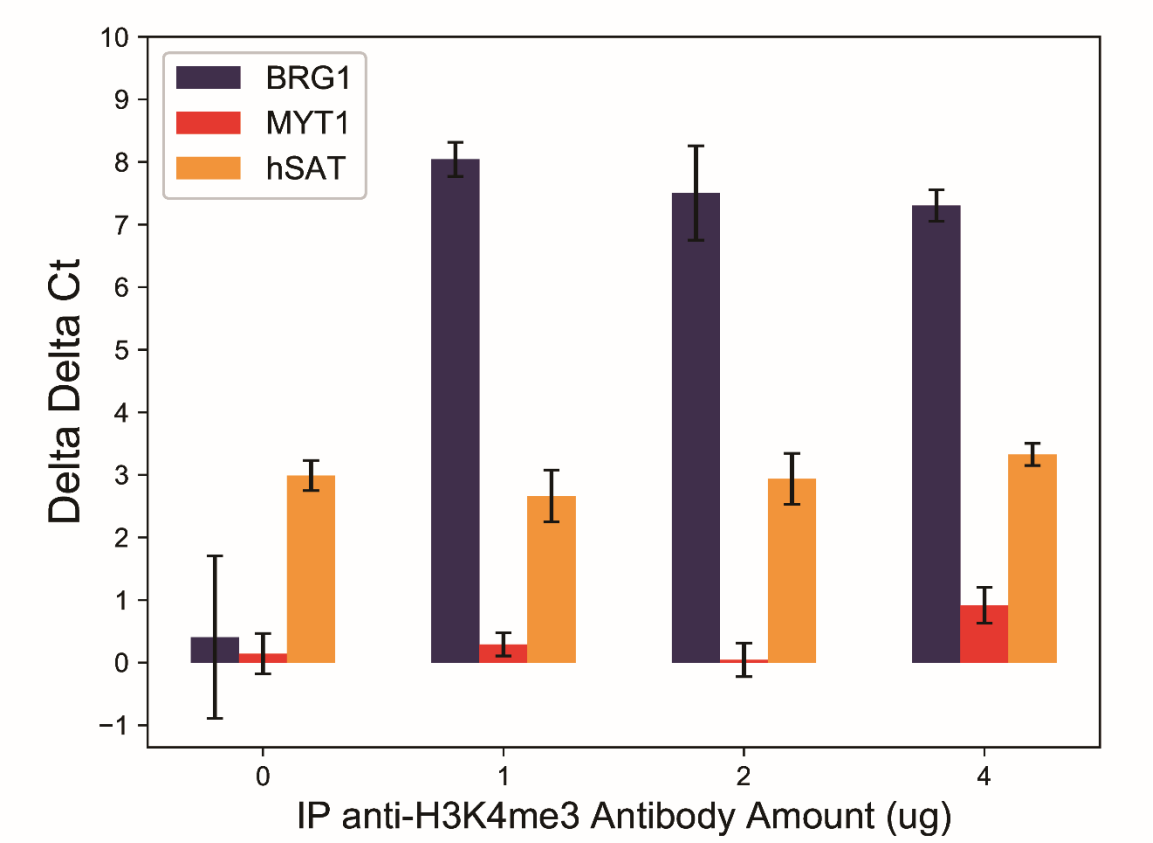

Figure 5-3 Average fold change results of ChIP using various amount of anti-H3K4me3 antibody per reaction. Error bars were standard deviations from three technical replicates under each condition. BRG1 stood for fold change of *Brg1_TSS* normalized to *C19_intergenic* in ChIP DNA compared to the input sample, MYT1 stood for fold change of *Myt1_TSS* normalized to *C19_intergenic*, hSAT stood for fold change of *SAT-alpha* normalized to *C19_intergenic*.

Additionally, in the case of the project to be developed in this chapter, one additional aspect needed confirmation: Whether a separately prepared fragmented chromatin can be used as input sample (100% input), instead of taking parts from the fragmented chromatin added to antibody-bead conjugates. After all, most traditional ChIP protocols take input samples for future analysis from chromatin prepared in bulk before aliquoting for ChIP.[75] Therefore, ChIP reactions were carried out in three formats to evaluate the concern discussed above. The antibody used here targeted H3K27me3, which was known to be a repressive mark, thus only the locus *MYT1_TSS*

should be associated with this mark as the related gene encodes a protein not expressed in the cell line. Format 1 resuspended antibody-functionalized beads in the combined lysis and digestion buffer and followed the protocol "ChIP in tube mimicking on-chip conditions" described in the section of "Chromatin immunoprecipitation (ChIP) under multiple conditions" in "Materials and Methods" for the rest of the experiment. Format 2 followed the protocol "ChIP in tube mimicking on-chip conditions" exactly. Format 3 followed the protocol "traditional ChIP" exactly contributing two 1% input samples. Duplicates were tested for each condition and the results were shown in **Figure 5-4**. The y axis was the fold change calculated based on 1% input taken from format 3 replicate 1 (F3_1_1%inp). Notice that the fold changes of three gene loci of the 100% input sample and the other 1% input sample were all close to 1, demonstrating that the chromatin fragmentation was consistent across replicates and parts of replicates, showing that simultaneously lysed and digested native chromatin suspension was homogenized well and that taking 1% out of 100% or taking 100% directly would not disturb the normalizing base line for qPCR. This result, along with the digestion patterns of all three input samples shown in **Figure B-8 to 10**, really strengthened the fidelity of the qPCR data in Figure 5-3. Furthermore, the 2nd replicate of format 2 and the 1st replicate of format 3 had no convincing C_t values thus no fold change values available due to little copies of Brg1_TSS in the samples, indicating that the antibodies targeted H3K27me3 with high specificity, showing little enrichment of gene loci not associated with the targeted histone modification. Both format 2 and 3 but not format 1 showed effective enrichment of *Myt1_TSS* with fold change similar to or over 5, validating that the corresponding two ChIP protocols were effective. One thing to notice is that both replicates in format 3 had fold change of SAT-alpha close to 1, which was expected as this gene loci was not supposed to be associated with either H3K4me3 or H3K27me3, but only H3K9me3. Such low enrichment of this gene locus was

not obtained in ChIP carried out in the conditions mimicking on-chip version using either antibody. This perhaps indicated that a bigger chromatin volume contributed to even less non-specific binding of any chromatin fragments to the antibody-bead conjugates as sufficient mixing/disturbing and enough dissipating volume existed when using the same volume and number of beads. This corresponded with previous literature[76] and the protocol of macroscale ChIP carried out by my collaborator at Mayo Clinic. However, as enlarging the fragmented chromatin's volume on chip would not be applicable for now, the results here confirming that the small-volume fragmented chromatin still qualifies for ChIP were exciting for the continuous development of the project.

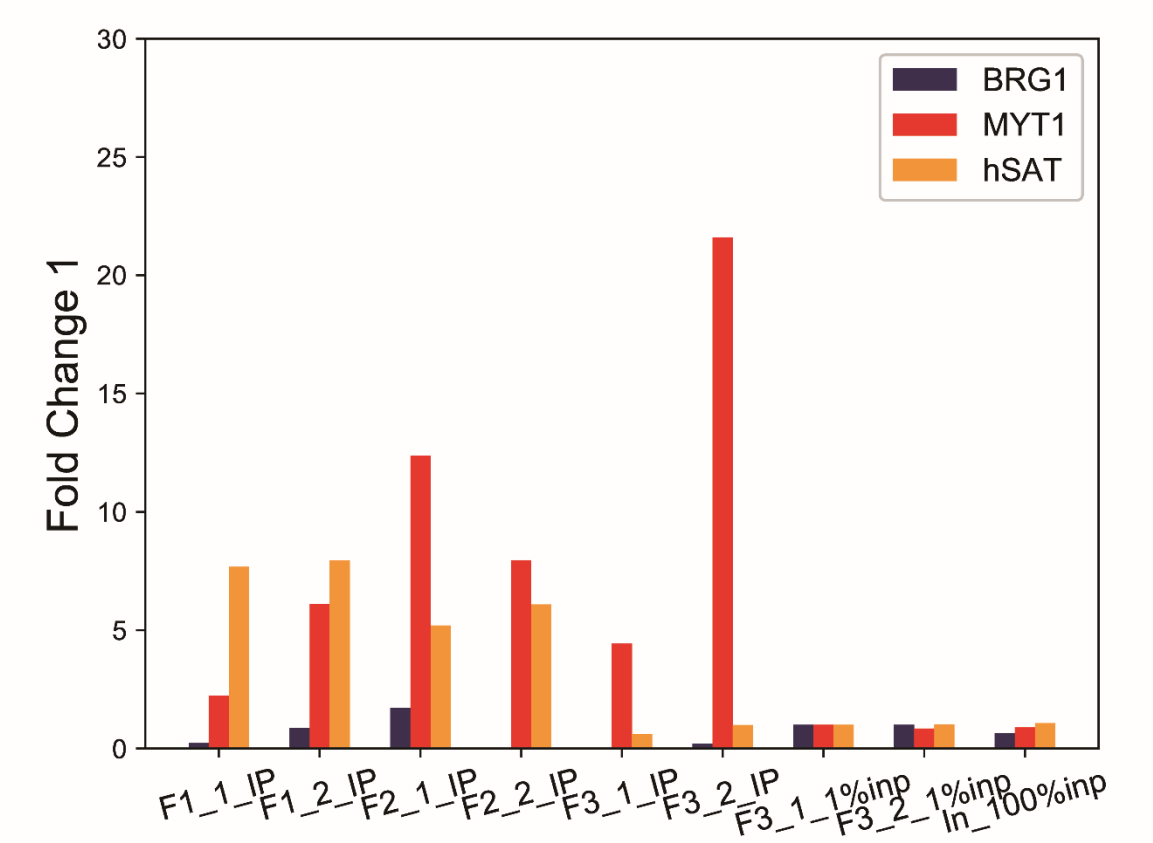

Figure 5-4 Fold change results of ChIP under three formats using 2-µg anti-H3K27me3 antibody per reaction. BRG1 stood for fold change of *Brg1_TSS* normalized to *C19_intergenic* in ChIP DNA compared to the input sample F3_1_1%inp, MYT1 stood for fold change of *Myt1_TSS* normalized to *C19_intergenic*, hSAT stood for fold change of *SAT-alpha* normalized to *C19_intergenic*.

Further comparison between 100% input samples prepared on chip (using the same device shown in Figure 5-1 with 2×STOP/ChIP buffer without beads) and in tube validated the consistency of chromatin fragmentation to a wider extent. As shown in **Figure 5-5**, the fold change of three gene loci in both in tube 100% input samples were all close to 1 using on-chip 100% input as the reference sample, though the on-chip ChIP did not work efficiently. This proved that for the ease of future development, input samples could be generated simply in tube without the need to operate another device if materials are limited and the qPCR results would still be convincing;

however, if sample consistency was emphasized, on-chip prepared 100% input would yield similarly qualified results as well.

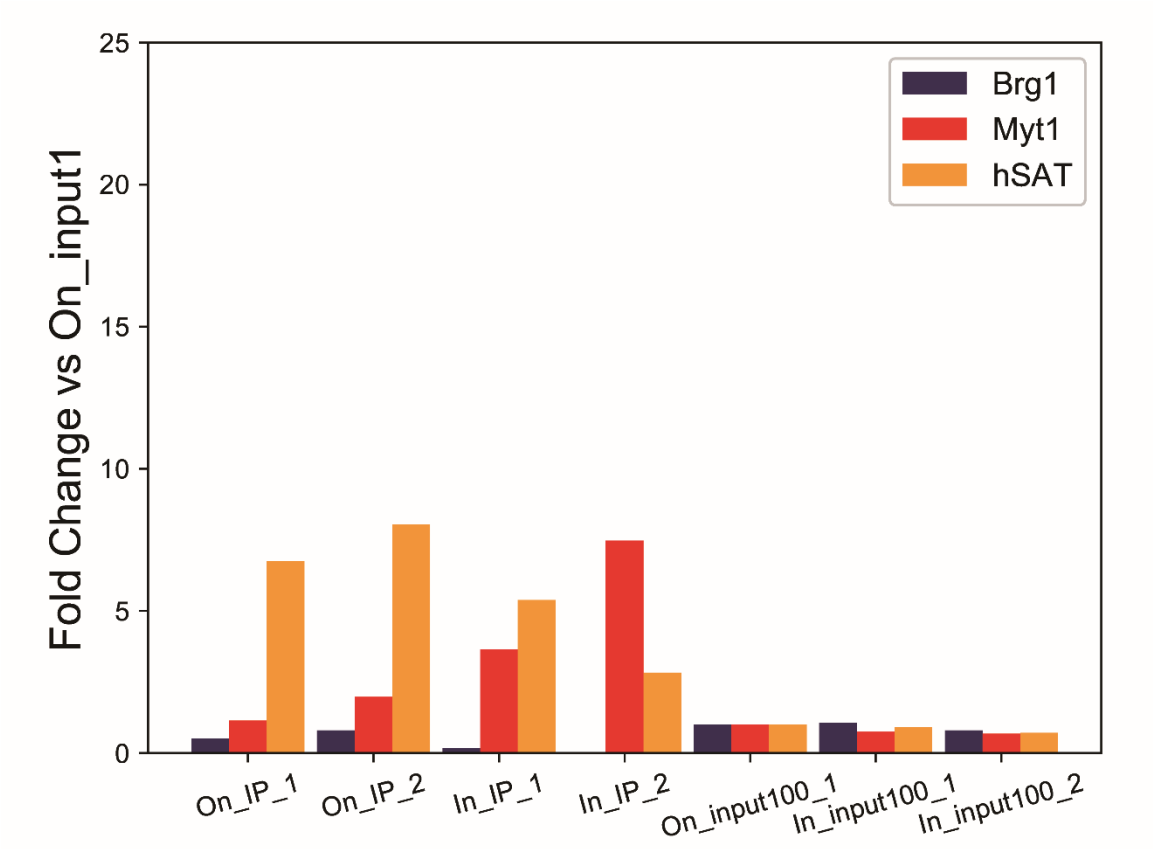

Figure 5-5 Fold change results of ChIP using 2-µg anti-H3K27me3 antibody per reaction. BRG1 stood for fold change of *Brg1_TSS* normalized to *C19_intergenic* in ChIP DNA compared to the input sample F3_1_1%inp, MYT1 stood for fold change of *Myt1_TSS* normalized to *C19_intergenic*, hSAT stood for fold change of *SAT-alpha* normalized to *C19_intergenic*. Each condition (on chip and in tube) had duplicates.

After the in-tube validation mimicking on-chip conditions obtaining qualified ChIP results using small volumes of fragmented chromatin and 100% input samples (prepared in tube), the characterization of optimal ratio among antibody, beads, and chromatin amount, and the characterization of immunoprecipitation incubation time duration (results shown in **Figure B-7**), ChIP on chip with manual washing was tested using 2-µg anti-H3K27me3 antibody per reaction

following “ChIP on chip” described in the section of “Chromatin immunoprecipitation (ChIP) under multiple conditions” in “Materials and Methods”. As shown in **Figure 5-6**, on-chip ChIP (orange) did not enrich the associated gene locus *Myt1_TSS* as effectively as in-tube ChIP mimicking on-chip conditions (green), while the non-associated *SAT-alpha* fold change spread out wider and averagely higher. However, the fact that *Brg1_TSS* was not enriched by on-chip ChIP was promising, indicating the (partial) existence of ChIP specificity on chip.

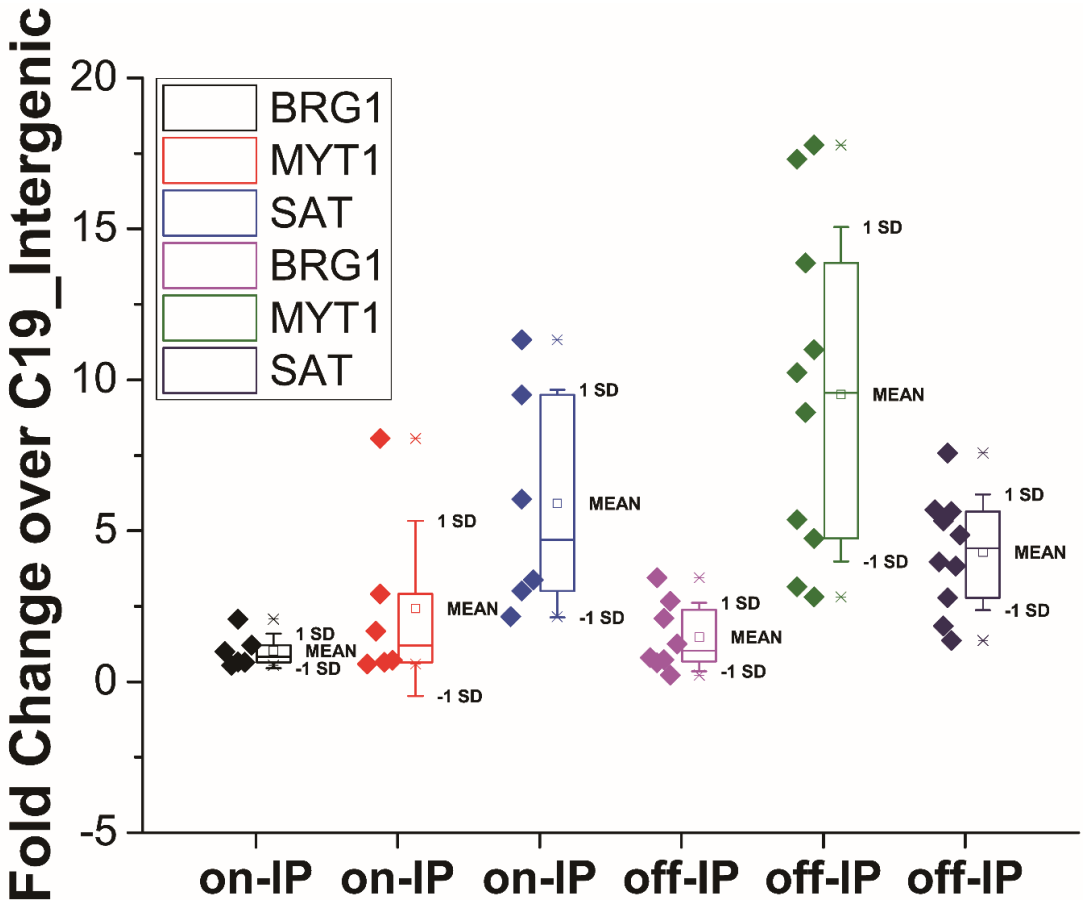

Figure 5-6 Fold change results of ChIP on chip (on-IP) and ChIP in tube mimicking on-chip conditions (off-IP). BRG1 stood for fold change of *Brg1_TSS* normalized to *C19_intergenic* in ChIP DNA compared to the input samples, MYT1 stood for fold change of *Myt1_TSS* normalized to *C19_intergenic*, SAT stood for fold change of *SAT-alpha* normalized to *C19_intergenic*. Input samples were in tube 100% input. On-chip ChIP had 6 replicates, ChIP in tube mimicking on-chip conditions had 10 replicates.

During the experiments, problems of on-chip ChIP were noticed and need to be solved to improve the enrichment efficiency. First, the beads resuspended in 2×STOP/ChIP buffer did not remain suspended or homogenized during the entire ChIP process (**Figure 5-7a**, no beads coming out at that specific moment). Beads precipitated easily during processing and aggregated with each other (**Figure 5-7b and c**). This led to the fact that the droplets injected with beads were not injected with same numbers of beads, and many droplets would not be injected with any antibody-bead conjugates. This means that part of the chromatin was not immunoprecipitated at all and thus sample bias happened during on-chip ChIP, as well as potential deviation from the optimal ratio among antibody, beads, and chromatin amount. Second, not all droplets contain cells, but the injection had to remain all the time. This caused waste of antibody-bead conjugates where antibody can be expensive and precious. Third, after collecting of beads containing droplets, whether for manual wash or after device wash for captured DNA elution directly, some magnetic beads were trapped in the oil phase and hard to be recovered. Bead loss contributed to the uncertainty of enrichment evaluation: does the better or worse fold change come from the fact that less ChIP DNA were eluted due to fewer beads, so it is artefact?

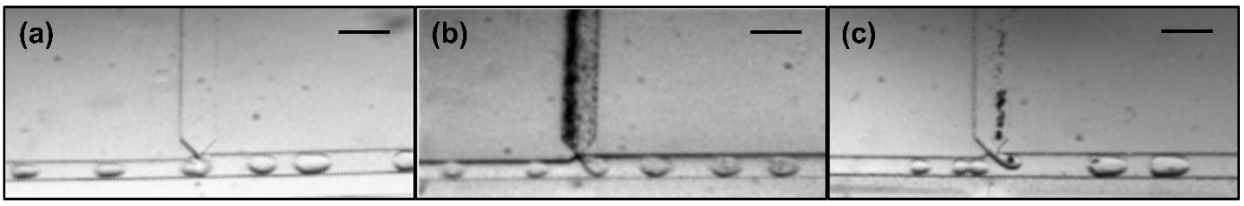

Figure 5-7 Illustration of bead containing 2×STOP/ChIP buffer injection problems. (a) No bead coming out. (b) Many beads coming out but on one edge. (c) Aggregated beads coming out as individual groups. Scale bars represent 200 μm.

Therefore, a strategy to quantitatively load antibody functionalized beads that remain suspended needs to be developed in order to reach a quantitative ChIP condition. Controlling bead

loss is also a significant aspect to improve on-chip ChIP. Furthermore, if a sorting mechanism can be integrated to the workflow teasing out the empty droplets without cells after the improvement mentioned above, expensive reagents such as antibody can be saved more. It is also more feasible controlling antibody-bead amount per cell or per two cells if the droplets are sorted and the percentage of single-cell and double-cell containing droplets are known. When these aspects are all improved, on-chip ChIP can be quantitatively evaluated which is more accurate than what has been obtained so far, and based on the in-tube ChIP results, it is promising that on-chip ChIP will work.

Conclusion

In this chapter, the characterization, optimization, and validation of droplet microfluidic ChIP was described in detail. With the existing data, two important aspects of the new ChIP format were validated demonstrating the possibility of doing ChIP in droplet microfluidic format. First, simultaneous lysis and digestion happening on the MNase processing device generated qualified fragmented chromatin for ChIP reaction with majority being mononucleosomes. Second, the ChIP conditions happening on the MNase processing device (adding antibody-bead conjugates to the chromatin containing droplets with quenching buffer directly without aliquoting 1% input as traditional ChIP does but using separately prepared 100% input) allowed convincing qPCR results and thus successful ChIP to enrich appropriate regions of chromatin. Though preliminary on-chip ChIP did not concentrate targeted regions significantly, the problems existing in the current device systems were pointed out and the potential solutions planned. Future work needs to be done as planned to fully develop droplet microfluidic ChIP.

References

1. H. Kimura, *Journal of human genetics*, 2013, **58**, 439-445.
2. D. Corujo and M. Buschbeck, *Cancers*, 2018, **10**.
3. A. J. Bannister and T. Kouzarides, *Cell Res*, 2011, **21**, 381-395.
4. Y. Obata, Y. Furusawa and K. Hase, *Immunology and cell biology*, 2015, **93**, 226-232.
5. Z. Chen, S. Li, S. Subramaniam, J. Y. Shyy and S. Chien, *Annu Rev Biomed Eng*, 2017, **19**, 195-219.
6. E. S. Chrun, F. Modolo and F. I. Daniel, *Pathology, research and practice*, 2017, **213**, 1329-1339.
7. M. Vermeulen, H. C. Eberl, F. Matarese, H. Marks, S. Denissov, F. Butter, K. K. Lee, J. V. Olsen, A. A. Hyman, H. G. Stunnenberg and M. Mann, *Cell*, 2010, **142**, 967-980.
8. J. Li, X. Xing, X. Zhang, B. Liang, Z. He, C. Gao, S. Wang, F. Wang, H. Zhang, S. Zeng, J. Fan, L. Chen, Z. Zhang, B. Zhang, C. Liu, Q. Wang, W. Lin, G. Dong, H. Tang, W. Chen, Y. Xiao and D. Li, *Environ Pollut*, 2018, **234**, 127-135.
9. F. S. Howe, H. Fischl, S. C. Murray and J. Mellor, *BioEssays : news and reviews in molecular, cellular and developmental biology*, 2017, **39**, 1-12.
10. C. Wang, J. E. Lee, B. Lai, T. S. Macfarlan, S. Xu, L. Zhuang, C. Liu, W. Peng and K. Ge, *Proceedings of the National Academy of Sciences of the United States of America*, 2016, **113**, 11871-11876.
11. W. A. Whyte, S. Bilodeau, D. A. Orlando, H. A. Hoke, G. M. Frampton, C. T. Foster, S. M. Cowley and R. A. Young, *Nature*, 2012, **482**, 221-225.
12. G. Ding, W. Li, J. Liu, Y. Zeng, C. Mao, Y. Kang and J. Shang, *Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie*, 2017, **94**, 326-331.
13. A. Rada-Iglesias, R. Bajpai, T. Swigut, S. A. Brugmann, R. A. Flynn and J. Wysocka, *Nature*, 2011, **470**, 279-283.
14. X. Yang, B. Hu, Y. Hou, Y. Qiao, R. Wang, Y. Chen, Y. Qian, S. Feng, J. Chen, C. Liu, G. Peng, F. Tang and N. Jing, *Cell Res*, 2018, **28**, 593-596.
15. Y. Qu, Q. Yang, J. Liu, B. Shi, M. Ji, G. Li and P. Hou, *Theranostics*, 2017, **7**, 2092-2107.
16. S. Li, S. Ali, X. Duan, S. Liu, J. Du, C. Liu, H. Dai, M. Zhou, L. Zhou, L. Yang, P. Chu, L. Li, R. Bhatia, D. E. Schones, X. Wu, H. Xu, Y. Hua, Z. Guo, Y. Yang, L. Zheng and B. Shen, *Cell reports*, 2018, **23**, 389-403.
17. R. Z. Jurkowska, S. Qin, G. Kungulovski, W. Tempel, Y. Liu, P. Bashtrykov, J. Stiefelmaier, T. P. Jurkowski, S. Kudithipudi, S. Weirich, R. Tamas, H. Wu, L. Dombrovski, P. Loppnau, R. Reinhardt, J. Min and A. Jeltsch, *Nature communications*, 2017, **8**, 2057.
18. H. Takeshima, M. Wakabayashi, N. Hattori, S. Yamashita and T. Ushijima, *Cancer Research*, 2014, **74**, 2314-2314.
19. Y. Atlasi and H. G. Stunnenberg, *Nature reviews. Genetics*, 2017, **18**, 643-658.
20. D. S. Johnson, A. Mortazavi, R. M. Myers and B. Wold, *Science*, 2007, **316**, 1497-1502.
21. R. Jothi, S. Cuddapah, A. Barski, K. Cui and K. Zhao, *Nucleic acids research*, 2008, **36**, 5221-5231.
22. G. Robertson, M. Hirst, M. Bainbridge, M. Bilenky, Y. Zhao, T. Zeng, G. Euskirchen, B. Bernier, R. Varhol, A. Delaney, N. Thiessen, O. L. Griffith, A. He, M. Marra, M. Snyder and S. Jones, *Nature methods*, 2007, **4**, 651-657.

23. R. Blecher-Gonen, Z. Barnett-Itzhaki, D. Jaitin, D. Amann-Zalcenstein, D. Lara-Astiaso and I. Amit, *Nature protocols*, 2013, **8**, 539-554.
24. G. K. Marinov, *Methods Mol Biol*, 2017, **1543**, 3-18.
25. I. Cheung, H. P. Shulha, Y. Jiang, A. Matevossian, J. Wang, Z. Weng and S. Akbarian, *Proceedings of the National Academy of Sciences of the United States of America*, 2010, **107**, 8824-8829.
26. Y. Kitagawa, N. Ohkura, Y. Kidani, A. Vandenbon, K. Hirota, R. Kawakami, K. Yasuda, D. Motooka, S. Nakamura, M. Kondo, I. Taniuchi, T. Kohwi-Shigematsu and S. Sakaguchi, *Nature immunology*, 2016, DOI: 10.1038/ni.3646.
27. R. Mathur, B. H. Alver, A. K. San Roman, B. G. Wilson, X. Wang, A. T. Agoston, P. J. Park, R. A. Shivdasani and C. W. Roberts, *Nature genetics*, 2016, DOI: 10.1038/ng.3744.
28. R. Hansel-Hertsch, D. Beraldi, S. V. Lensing, G. Marsico, K. Zyner, A. Parry, M. Di Antonio, J. Pike, H. Kimura, M. Narita, D. Tannahill and S. Balasubramanian, *Nature genetics*, 2016, **48**, 1267-1272.
29. S. Beyaz, J. H. Kim, L. Pinello, M. E. Xifaras, Y. Hu, J. Huang, M. A. Kerenyi, P. P. Das, R. A. Barnitz, A. Herault, R. Dogum, W. N. Haining, O. H. Yilmaz, E. Passegue, G. C. Yuan, S. H. Orkin and F. Winau, *Nature immunology*, 2016, DOI: 10.1038/ni.3644.
30. M. de Dieuleveult, K. Yen, I. Hmitou, A. Depaux, F. Boussouar, D. Bou Dargham, S. Jounier, H. Humbertclaude, F. Ribierre, C. Baulard, N. P. Farrell, B. Park, C. Keime, L. Carriere, S. Berlivet, M. Gut, I. Gut, M. Werner, J. F. Deleuze, R. Olaso, J. C. Aude, S. Chantalat, B. F. Pugh and M. Gerard, *Nature*, 2016, **530**, 113-116.
31. A. Barski, S. Cuddapah, K. Cui, T.-Y. Roh, D. E. Schones, Z. Wang, G. Wei, I. Chepelev and K. Zhao, *Cell*, 2007, **129**, 823-837.
32. H. Yan, S. Tian, S. L. Slager, Z. Sun and T. Ordog, *American journal of epidemiology*, 2016, **183**, 96-109.
33. A. R. Wu, J. B. Hiatt, R. Lu, J. L. Attema, N. A. Lobo, I. L. Weissman, M. F. Clarke and S. R. Quake, *Lab on a chip*, 2009, **9**, 1365-1370.
34. A. R. Wu and S. R. Quake, *Cold Spring Harbor protocols*, 2016, **2016**, pdb prot084996.
35. A. R. Wu, T. L. Kawahara, N. A. Rapicavoli, J. van Riggelen, E. H. Shroff, L. Xu, D. W. Felsher, H. Y. Chang and S. R. Quake, *Lab on a chip*, 2012, **12**, 2190-2198.
36. Z. Cao, C. Chen, B. He, K. Tan and C. Lu, *Nature methods*, 2015, **12**, 959-962.
37. Z. Cao and C. Lu, *Analytical chemistry*, 2016, **88**, 1965-1972.
38. T. W. Murphy, Y. P. Hsieh, S. Ma, Y. Zhu and C. Lu, *Analytical chemistry*, 2018, **90**, 7666-7674.
39. S. Y. Teh, R. Lin, L. H. Hung and A. P. Lee, *Lab on a chip*, 2008, **8**, 198-220.
40. R. Seemann, M. Brinkmann, T. Pfohl and S. Herminghaus, *Rep Prog Phys*, 2012, **75**, 016601.
41. A. Huebner, S. Sharma, M. Srisa-Art, F. Hollfelder, J. B. Edel and A. J. Demello, *Lab on a chip*, 2008, **8**, 1244-1254.
42. O. J. Dressler, R. M. Maceiczyk, S. I. Chang and A. J. deMello, *Journal of biomolecular screening*, 2014, **19**, 483-496.
43. C. N. Baroud, F. Gallaire and R. Dangla, *Lab on a chip*, 2010, **10**, 2032-2045.
44. G. F. Christopher and S. L. Anna, *Journal of Physics D: Applied Physics*, 2007, **40**, R319-R336.
45. M. Leman, F. Abouakil, A. D. Griffiths and P. Tabeling, *Lab on a chip*, 2015, **15**, 753-765.

46. M. Chabert and J. L. Viovy, *Proceedings of the National Academy of Sciences of the United States of America*, 2008, **105**, 3191-3196.
47. H. H. Gorris and D. R. Walt, *Angew Chem Int Ed Engl*, 2010, **49**, 3880-3895.
48. E. W. Kemna, R. M. Schoeman, F. Wolbers, I. Vermes, D. A. Weitz and A. van den Berg, *Lab on a chip*, 2012, **12**, 2881-2887.
49. D. J. Eastburn, A. Sciambi and A. R. Abate, *Analytical chemistry*, 2013, **85**, 8016-8021.
50. T. M. Tran, F. Lan, C. S. Thompson and A. R. Abate, *Journal of Physics D: Applied Physics*, 2013, **46**, 114004.
51. P. Garstecki, M. J. Fuerstman, H. A. Stone and G. M. Whitesides, *Lab on a chip*, 2006, **6**, 437-446.
52. H. N. Joensson and H. Andersson Svahn, *Angew Chem Int Ed Engl*, 2012, **51**, 12176-12192.
53. L. Mazutis, J. Gilbert, W. L. Ung, D. A. Weitz, A. D. Griffiths and J. A. Heyman, *Nature protocols*, 2013, **8**, 870-891.
54. T. P. Lagus and J. F. Edd, *Journal of Physics D: Applied Physics*, 2013, **46**, 114005.
55. A. Rakszewska, J. Tel, V. Chokkalingam and W. T. S. Huck, *NPG Asia Materials*, 2014, **6**, e133.
56. A. R. Abate, C. H. Chen, J. J. Agresti and D. A. Weitz, *Lab on a chip*, 2009, **9**, 2628-2631.
57. J. F. Edd, D. Di Carlo, K. J. Humphry, S. Koster, D. Irimia, D. A. Weitz and M. Toner, *Lab on a chip*, 2008, **8**, 1262-1264.
58. A. Rotem, O. Ram, N. Shoresh, R. A. Sperling, A. Goren, D. A. Weitz and B. E. Bernstein, *Nature biotechnology*, 2015, **33**, 1165-1172.
59. L. Mazutis and A. D. Griffiths, *Lab on a chip*, 2012, **12**, 1800-1806.
60. Y. Ding, X. C. i Solvas and A. deMello, *The Analyst*, 2015, **140**, 414-421.
61. A. Sciambi and A. R. Abate, *Lab on a chip*, 2015, **15**, 47-51.
62. V. Chokkalingam, Y. Ma, J. Thiele, W. Schalk, J. Tel and W. T. Huck, *Lab on a chip*, 2014, **14**, 2398-2402.
63. E. Brouzes, A. Carniol, T. Bakowski and H. H. Strey, *RSC advances*, 2014, **4**, 38542-38550.
64. Y.-C. Tan, Y. L. Ho and A. P. Lee, *Microfluidics and Nanofluidics*, 2006, **3**, 495-499.
65. D. R. Link, S. L. Anna, D. A. Weitz and H. A. Stone, *Physical Review Letters*, 2004, **92**.
66. B. O'Donovan, T. Tran, A. Sciambi and A. Abate, *Journal of visualized experiments : JoVE*, 2014, DOI: 10.3791/50913.
67. L. M. Fidalgo, C. Abell and W. T. Huck, *Lab on a chip*, 2007, **7**, 984-986.
68. L. Li, J. Q. Boedicker and R. F. Ismagilov, *Analytical chemistry*, 2007, **79**, 2756-2761.
69. A. R. Abate, T. Hung, P. Mary, J. J. Agresti and D. A. Weitz, *Proceedings of the National Academy of Sciences of the United States of America*, 2010, **107**, 19163-19166.
70. P. Tabeling, *Lab on a chip*, 2009, **9**, 2428-2436.
71. S. Ma, J. M. Sherwood, W. T. Huck and S. Balabani, *Lab on a chip*, 2014, **14**, 3611-3620.
72. J. D. Tice, H. Song, A. D. Lyon and R. F. Ismagilov, *Langmuir*, 2003, **19**, 9127-9133.
73. T. H. Ho, R. N. Nateras, H. Yan, J. G. Park, S. Jensen, C. Borges, J. H. Lee, M. D. Champion, R. Tibes, A. H. Bryce, E. M. Carballido, M. A. Todd, R. W. Joseph, W. W. Wong, A. S. Parker, M. L. Stanton and E. P. Castle, *PloS one*, 2015, **10**, e0132831.
74. M. F. Carey, C. L. Peterson and S. T. Smale, *Cold Spring Harbor protocols*, 2009, **4**, doi:10.1101/pdb.prot5279.
75. T. I. Lee, S. E. Johnstone and R. A. Young, *Nature protocols*, 2006, **1**, 729-748.
76. V. Spencer, *Methods*, 2003, **31**, 67-75.

Chapter 6 Conclusions and Future Work

The previous chapters described automated sample processing for three different epigenomic assays (MNase-seq, MBA-seq, and ChIP-seq) using droplet microfluidics. The feasibility of these developed methodology has been characterized and validated. More needs to be completed to further improve these platforms' performance and verify their applicability in clinical settings in order to utilize these techniques for epigenomic diagnostics and therapeutics eventually.

Nucleosome Preparation from Clinical Samples for MNase-seq

Chapter 3 validated the MNase-processing device's efficacy in preparing qualified samples for MNase-seq with immortal cell line cells. Whether this strategy works with clinical samples need to be confirmed for practical considerations. Preliminary data of patient studies using this droplet microfluidic platform has been collected by processing buffy coat. Buffy coat is a representative of clinically-relevant patient sample as it is a major component of whole blood without red blood cells or plasma.[1] The buffy coat collection method for this thesis was described in Chapter 4. Flow rates, MNase concentrations, and cell suspension conditions were the same as the optimal ones in Chapter 3. The selected disease model was septic shock, a subset of sepsis where after causing organ dysfunctions the widespread, dysregulated host responses to infection are drastically severe to induce increased risk of mortality.[2] Sepsis and septic shock have been recognized as global health issues and remained as one of the leading causes of in-hospital deaths in the USA (between 33% and 50%).[2] Moreover, sepsis and septic shock have complicated gene

expression regulating networks during the progress, thus increasing efforts are being made to further decipher the disease so that reliable diagnostic marks and therapeutic targets can be discovered to offer an improved, urgent, and correct clinical care of septic patients.[3] With the goal to discover the underlying nucleosome positioning mechanisms in septic shock, the buffy coat samples from septic shock patients, matched non-septic shock patients, and healthy controls were processed with the platform developed in Chapter 3. **Figure 6-1** confirmed the successful droplet microfluidic processing of buffy coat samples extracted from whole blood from five pairs of matched septic and non-septic shock patients, and three healthy controls (duplicates each patient, thus twenty-six samples processed on chip or off chip/in tube, fifty-two samples processed in total). Mononucleosome yields of all samples were consistent with the characterization results generated from Jurkat cells in Chapter 3 at around 80%, not significantly different from those off-chip controls mimicking the on-chip conditions, indicating no bias was induced by the device but only the biology of MNase contributed to the success of sample preparation. These samples were also sequenced and bioinformatic analysis is undergoing. Preliminary results from part of the patient pairs were shown as follows.

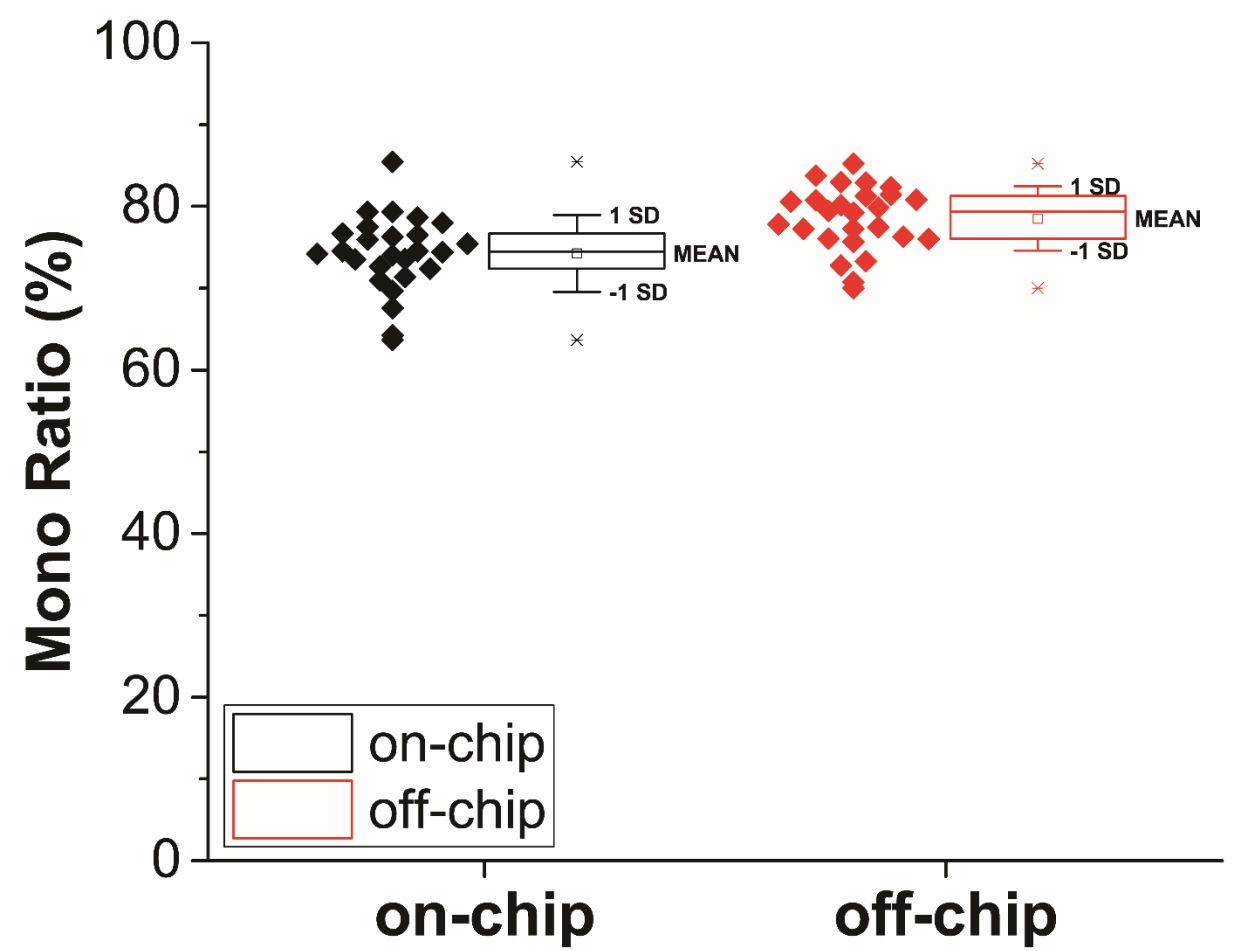

Figure 6-1 Mononucleosome yields of buffy coat samples processed on the device developed in Chapter 3 and in tube as controls. For each group of data, scattered dots of all data were on the left, and box plots with mean and one standard deviation as whiskers were on the right. Each data point represented the result of one sample containing 125,000 cells in 30-µL density-adjusted PBS.

Figure 6-2 showed the counted nucleosomes from libraries prepared from buffy coat mononucleosomal DNA processed on the device ("on") and in the tube mimicking on-chip conditions ("in"). Like the results shown in Figure 3-16, similar numbers of nucleosomes were counted in all samples, proving little bias existed in the droplet microfluidic platform. The absolute values of nucleosome peaks also fit with the theoretical number of nucleosome in human genome. This proved that clinical-relevant samples processed by the developed droplet microfluidic platform are capable of generating qualified sequencing results for further interpretation, which in

turn verified the potential of integrating this platform to clinical routine to assist the promotion of epigenetic diagnosis and therapy development.

Figure 6-3 showed the nucleosome positioning surrounding TSSs of all genes of the replicate samples from 0620A and B patients. Again, the high correlation of the peak and valleys of the nucleosome levels between on-chip duplicates and between on-chip and off-chip (in-tube) processed samples proved little bias was introduced by the microfluidic system and the samples were showing data that match with precedent literatures: well-positioned nucleosomes right up-stream and down-stream of TSSs and more fuzzier distribution going to the gene body.[4] Interestingly, the profile shapes of normalized nucleosome read density around TSSs of all genes were also similar between matched patients of septic shock and non-septic shock patients (**Figure 6-4**, from left to right were results of sample pairs collected on 0522, 0620, and 0628, respectively), while further IPA analysis (Ingenuity Pathway Analysis, QIAgen) showed that only a small number of genes (~100) of these two types of patients encountered dynamic nucleosomes within 2kb of TSSs including *FOXP1* (related to quiescence of CD4+ T lymphocytes), *FN1* (related to response to ischemia), and *FOS4* (related to inflammatory response such as cytokine signaling and multiple immune system signaling pathways) indicating that the accessibility to these genes' promoter regions were being regulated and changed comparing these two conditions. Therefore, perhaps only a few genes were functioning as the driving forces or results of the immune system's response to the widespread infection and inflammation in septic shock condition. If validated with a large-scale data set collected from more patients, this would be extraordinarily exciting from the aspect of elucidating the epigenetic regulating mechanisms involved in septic shock.

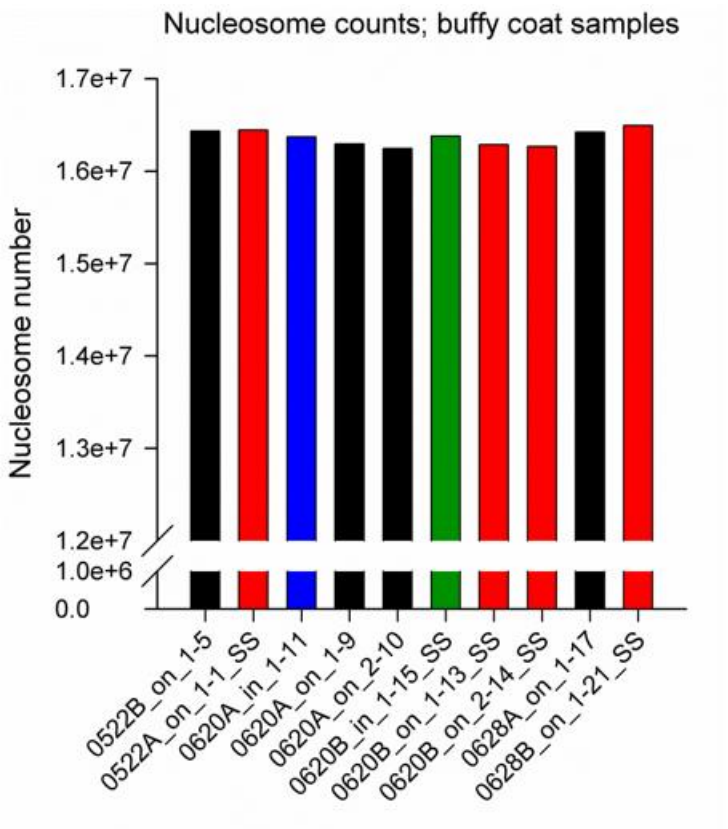

Figure 6-2 Nucleosomes counted in each library constructed from buffy coat mononucleosomal DNA prepared by the Chapter 3 device or in tube. The sample names were formatted as collected date followed by either "on" (on chip) or "in" (in tube, or off chip) with sample number and replicate number. The last number was library number when constructing them for sequencing but irrelevant to sample identity. "SS" stood for "septic shock".

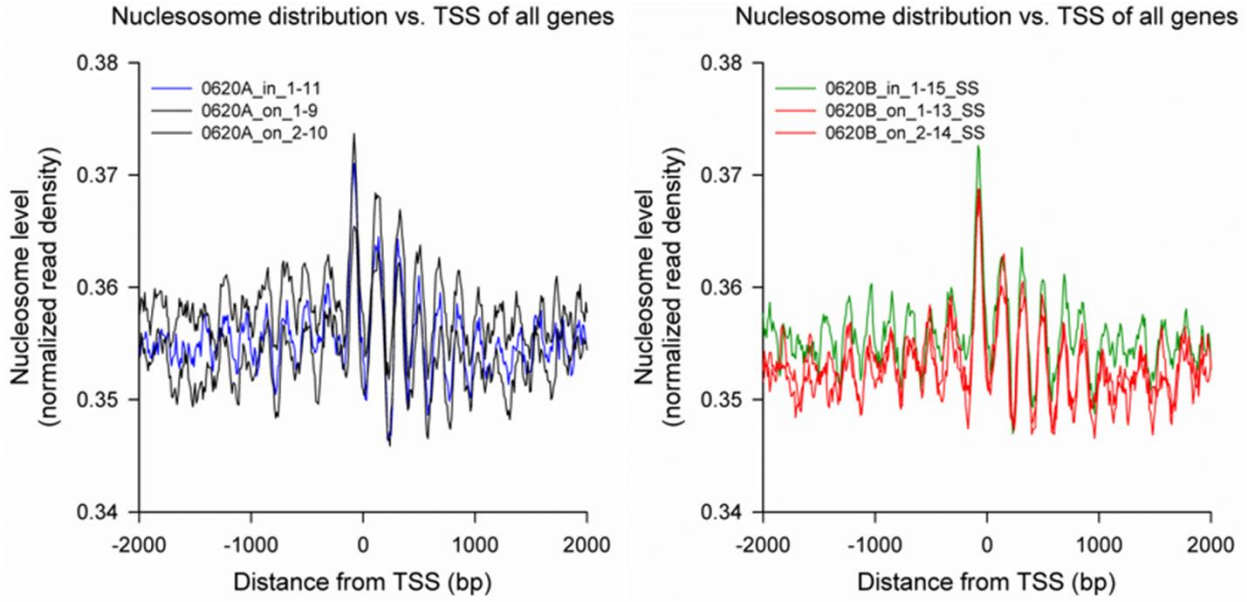

Figure 6-3 Normalized nucleosome read density surrounding TSSs of buffy coat extracted from whole blood of patients labeled as 0620A and 0620B, the latter being the septic shock patient. Notice the high correlation of the peak and valley of nucleosome levels among the one in-tube processed replicate and two on-chip processed replicates of each patient.

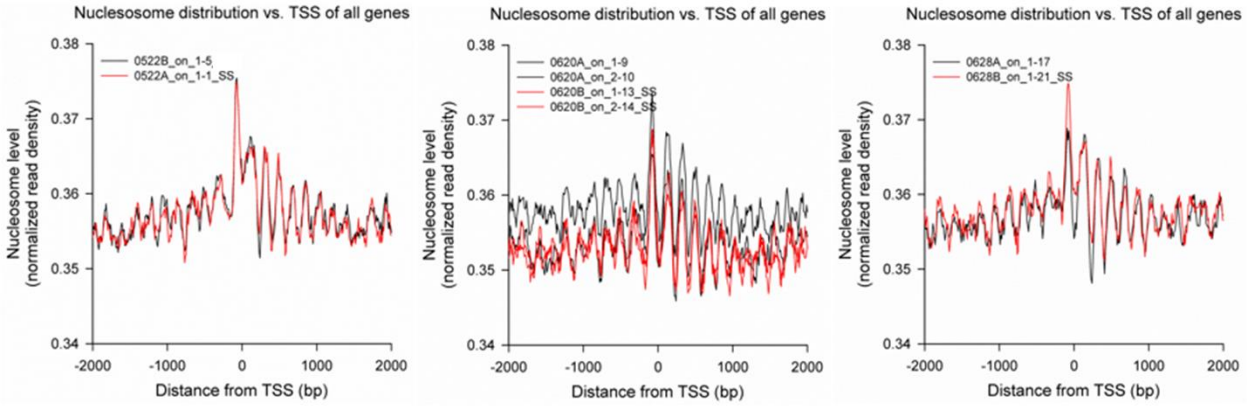

Figure 6-4 Normalized nucleosome read density surrounding TSSs of buffy coat extracted from whole blood of three pairs of patients. Notice the high correlation of the peak and valley of nucleosome levels between each matched pair.

These preliminary results showed promising future in elucidating epigenomic regulation changes in septic shock from the perspective of nucleosome positioning using the droplet microfluidic device developed in Chapter 3. Besides the completion of the bioinformatic analysis of samples from the other two pairs of patients and healthy controls, more patients need to be enrolled to collect "big data" to ensure the fidelity of the diagnosing insights summarized from the nucleosome mapping and identifying affected genes in septic shock. Moreover, septic shock can be the deadly result of the worsened sepsis development. Nucleosome positioning profiling along the time trajectory of sepsis patients' disease development can reveal more precious information of the mechanisms in epigenetic regulation of immune-related gene expression as sepsis develops so that involved epigenetic marks along with the regulated genes can be identified as prognostic marks in future septic diagnosis and treatment to predict the patients with higher risks of turning into septic shock in order to take better care of them. Therefore, processing the buffy coat from the same patients at different septic developing stages is highly expected as well to further the investigation. Furthermore, patients recovered from sepsis or septic shock may have changed epigenetic profiles casting long-term effects on their immune systems.[5] Thus, continuous sample collections and nucleosome positioning studies via tracking septic shock survivors for a certain duration of time should also be included in the long-term goal of this project to fully understand the epigenetic regulating mechanisms and related long-lasting impact involved in the process.

Additionally, the device developed in Chapter 3 also processed porcine buffy coat efficiently as shown in **Figure 6-5**. The porcine buffy coat was from a pig model introduced with sepsis. Whole blood samples were collected at 4 different time points and buffy coat extracted within 8 hours after the blood collected. The consistent native cell chromatin digestion efficiency showed that this developed platform has wide applicability in nucleosome positioning studies on

different subjects. Therefore, if time trajectory studies of human patients are challenging, animal model study with this platform can be a good alternative.

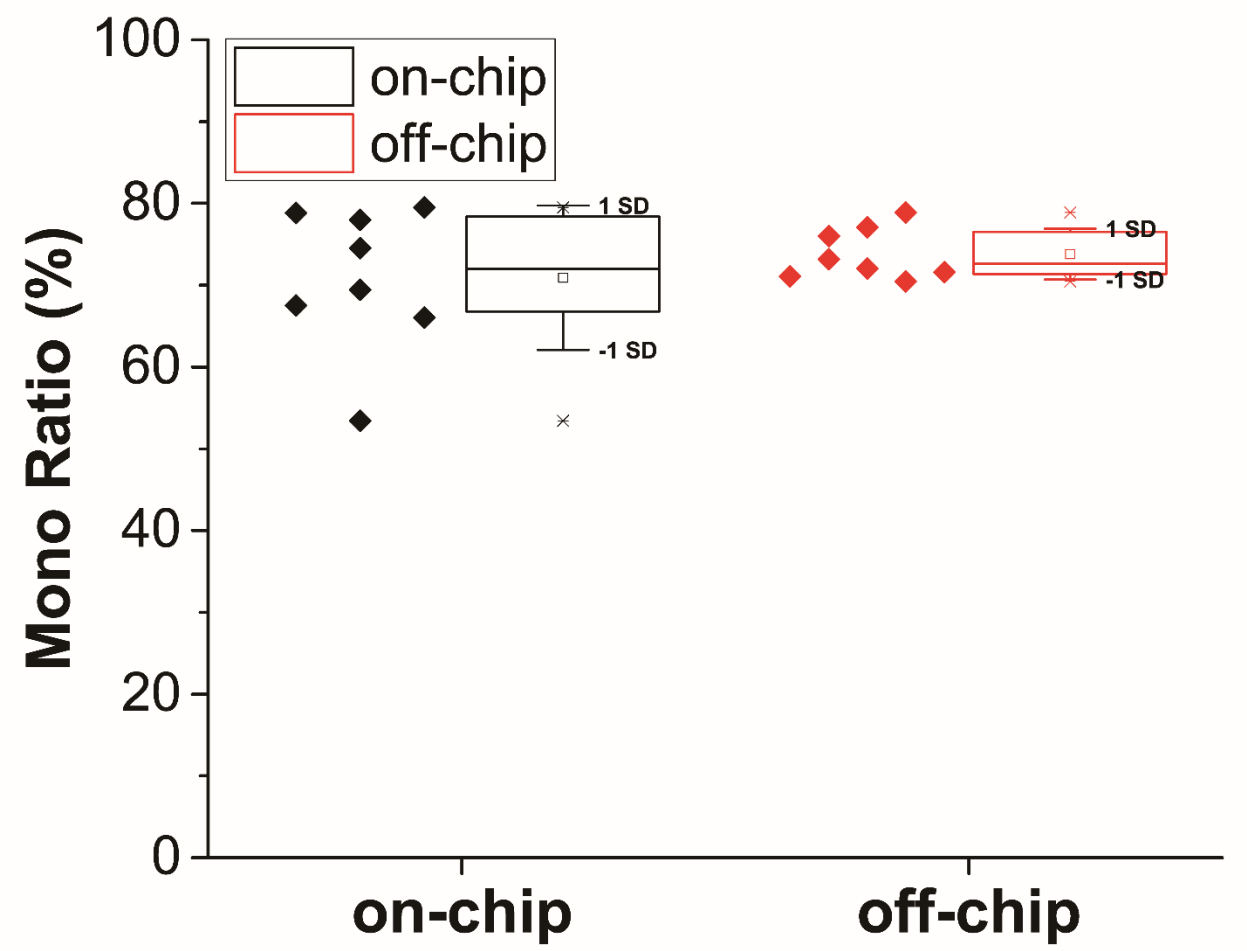

Figure 6-5 Mononucleosome yields of porcine buffy coat samples processed on the device developed in Chapter 3 and in tube as controls. For each group of data, scattered dots of all data were on the left, and box plots with mean and one standard deviation as whiskers were on the right. Each data point represented the result of one sample containing 166,666 cells in 40-µL density-adjusted PBS.

Further Optimization of Device Design to Collect Linker DNA for MBA-seq

Chapter 4 described a droplet microfluidic device that can provide two samples for two different epigenomic assays from the same batch of sample cells through passive aliquoting the generated droplets containing cells and MNase to two directions for different lengths of digestion time duration. Although the preliminary data proved success in enriching nucleosome-free linker DNA regions, more aspects can be improved.

First, active mechanisms to control sample aliquoting may be included to replace the current passive mechanism. Current mechanism simply relies on the bifurcating design to direct droplets to two directions. Thus, the ratio of droplets going to MBA side versus droplets going to MNase side is determined by the pressure drop difference between the two routes, which also determine the delay time on each side. This means anything that can disrupt the pressure drop in either branch such as a PDMS fiber or a piece of dust might affect the droplet aliquoting ratio and thus the DNA yield on each side. Therefore, if an active aliquoting mechanism such as valves can be integrated to the platform, quantitative percentage of the sample plug can be directed to either side, leading to a more accurate control of DNA amount on both sides.

Second, digestion for MBA-seq DNA can be controlled by lowering MNase concentration rather than the current strategy, which is shortening delay time. Although current strategy was proved to be working, it requires two branches of delay channels of different lengths and thicknesses in order to ensure the appropriately short digestion time for MBA DNA and long enough digestion time for MNase DNA. The device footprint was not compacted enough. **Figure 6-6** proved that 0.2×MNase (i.e., 4% of the current MNase concentration) digesting 50-µL, 5 million per mL cells for 210 s in tube generated qualified MBA DNA with significant enrichment of both positive loci tested by qPCR. This indicated the applicability of starting with low MNase

concentration to digest for MBA DNA and increasing the concentration to the current level to collect for MNase DNA with the same delay channel. In this situation, the device footprint could be reduced and additionally, the aliquoting ratio of samples between the two assays can be controlled by the timing increasing MNase concentration. Notice that even though shorter digestion time (<210 s) at lower concentration of MNase yielded more qualified MBA DNA with higher enrichment on both loci, it would not be different from the current design to have two different lengths of delay channels if adopting those conditions. Moreover, as long as the enrichment reaches the empirical threshold (fold change ≥ 5), the MBA DNA should be able to provide convincing sequencing results.

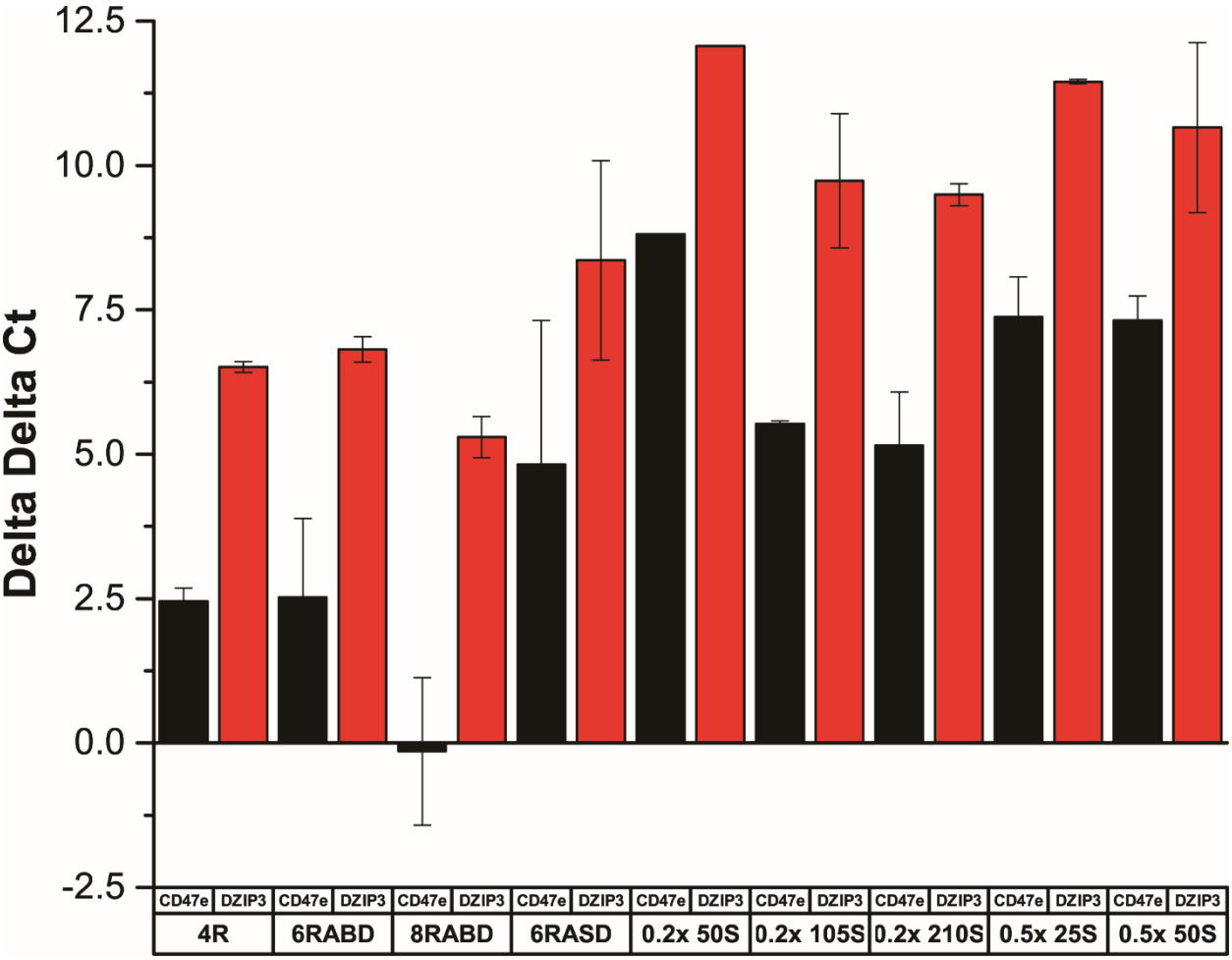

Figure 6-6 Fold change expressed in the format of ΔΔCt (i.e., $\log_2$(fold change)) of all devices tested in Chapter 4 as well as in-tube test with lower concentration of MNase and shorter digestion time. Notice the similar performance of samples processed by 0.2×MNase and digested for 210 s and samples processed by the optimal device 6RSAD with 5×MNase (i.e., the current MNase concentration).

Optimization and Automation of Droplet Microfluidic ChIP Workflow

As discussed in Chapter 5, current ChIP workflow needs more optimization to obtain efficient ChIP DNA from on-chip processing. Besides the necessity to optimize bead loading and droplet sorting, more parts of the entire ChIP workflow needs to be automated. **Figure 6-7** illustrated the proposed droplet ChIP workflow. Module 1 has been completed with slight modification (simultaneous lysis and digestion in Chapter 3). Module 2 is under development with

the efforts described in Chapter 5. Module 3 needs to be started as RNase A and Proteinase K treatment takes several hours to complete before DNA can be purified. Automating this step on droplet microfluidic platform will benefit not only ChIP assay but also MNase-seq assay and many other chromatin-based epigenomic assays. A proposed strategy is adapted from a cylindrical heater for real-time PCR[6] and the modified version is shown in **Figure 6-8**. The heater was fabricated by the electronic shop in the department of chemistry. Droplet collecting tubing was wrapped around the heater set at 65°C to activate RNase A and Proteinase K. Currently the strategy is to add these two enzymes simultaneously to the quenching buffer and inject to the droplets containing digested chromatin or antibody-bead conjugate bound chromatin after wash. Preliminary data showed the applicability of this strategy, and more experiments need to be done to optimize the performance and recover more DNA as final products.

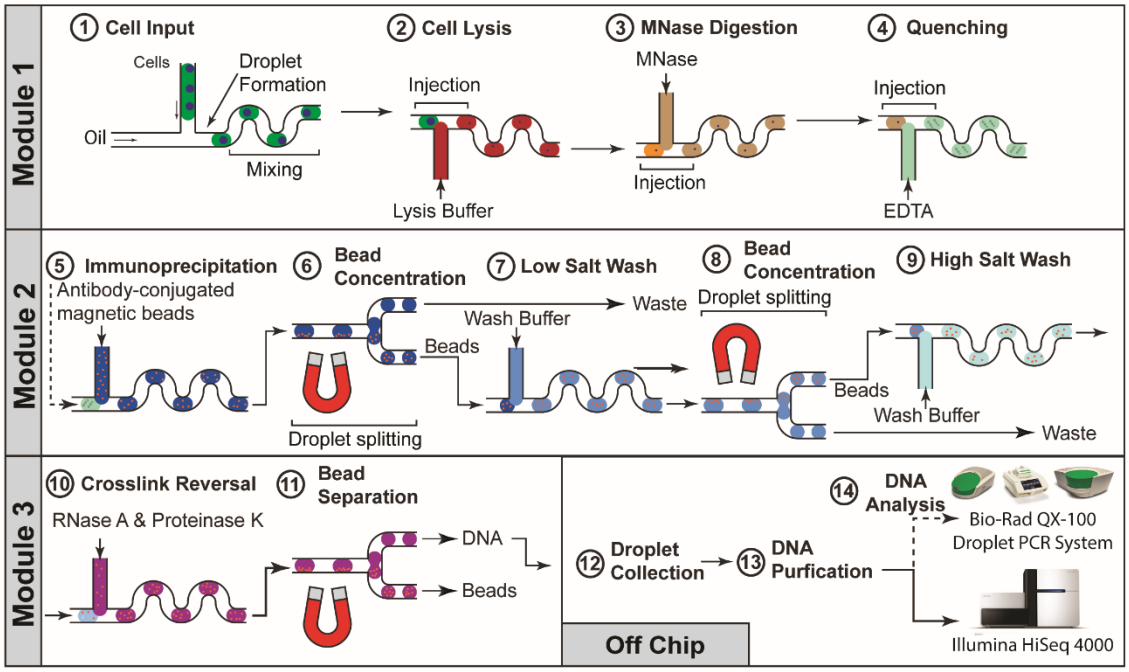

Figure 6-7 Proposed droplet ChIP workflow to be completed.

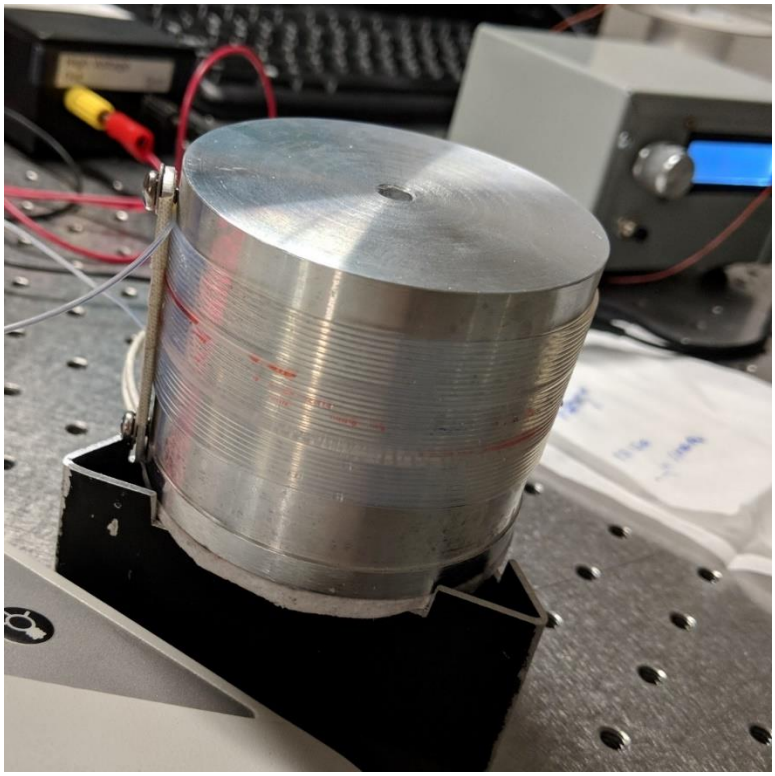

Figure 6-8 In-house fabricated cylindrical heater for integrated RNase A and Proteinase K treatment of chromatin using droplet microfluidics.

Moreover, currently the droplet ChIP is native ChIP (N-ChIP), processing cells without fixation. N-ChIP provides qualified data when the targets are histone modification marks, as the nucleosome structures are stable; however, if the targets are transiently bound proteins to functional DNA regions, it is better to fix the cells to preserve the binding event of such short-term binding proteins such as transcriptional factors through covalent bonds (X-ChIP).[7] Therefore, validating X-ChIP on the same droplet microfluidic platform will also be significant to broaden the application fields for the proposed droplet ChIP strategy.

References

1. S. B. Harris and C. D. Hillyer, 2007, DOI: 10.1016/b978-0-443-06981-9.50017-x, 183-204.
2. M. Cecconi, L. Evans, M. Levy and A. Rhodes, *The Lancet*, 2018, **392**, 75-87.
3. T. van der Poll, F. L. van de Veerdonk, B. P. Scicluna and M. G. Netea, *Nature reviews. Immunology*, 2017, **17**, 407-420.
4. C. Jiang and B. F. Pugh, *Nature reviews. Genetics*, 2009, **10**, 161-172.
5. F. Pinheiro da Silva and M. C. C. Machado, *Frontiers in immunology*, 2017, **8**, 1389.
6. A. C. Hatch, T. Ray, K. Lintecum and C. Youngbull, *Lab on a chip*, 2014, **14**, 562-568.
7. S. A. David, B. Piegu, C. Hennequet-Antier, M. Pannetier, T. Aguirre-Lavin, S. Crochet, T. Bordeau, N. Courousse, A. Brionne, Y. Bigot, A. Collin and V. Coustham, *Biological procedures online*, 2017, **19**, 10.

APPENDICES

Appendix A

Mask Designs

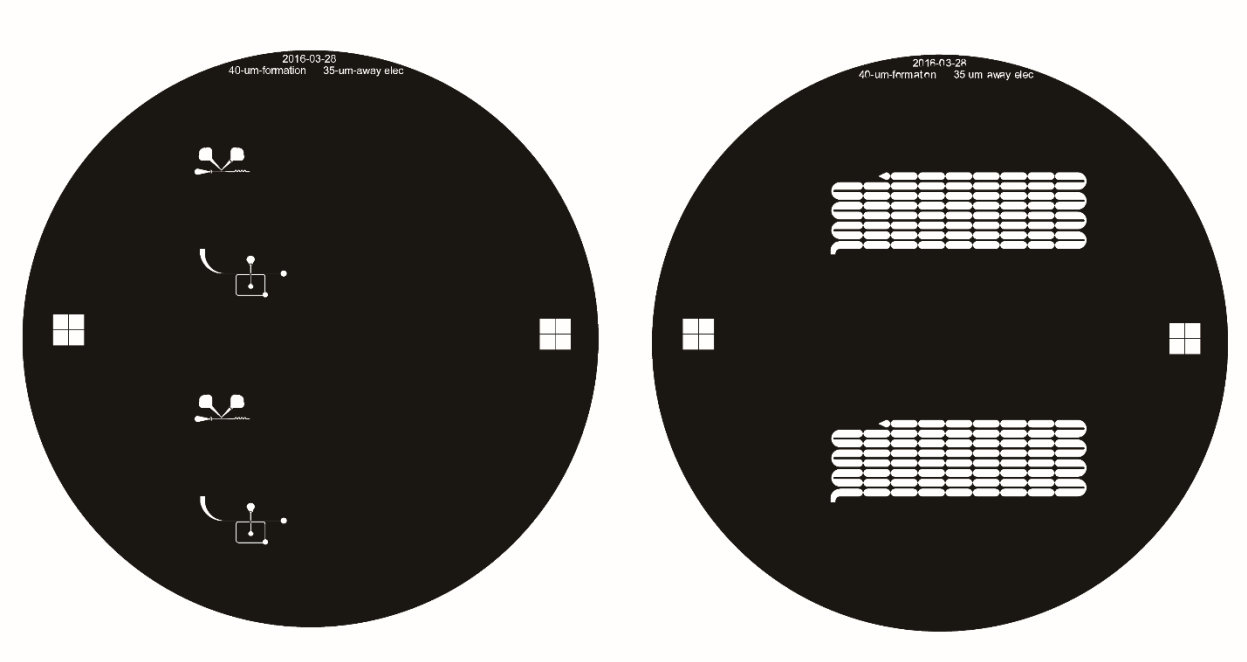

Figure A-1 Photomask of the MNase processing device design after translating AutoCAD files for printing by CAD/Art Services, Inc.

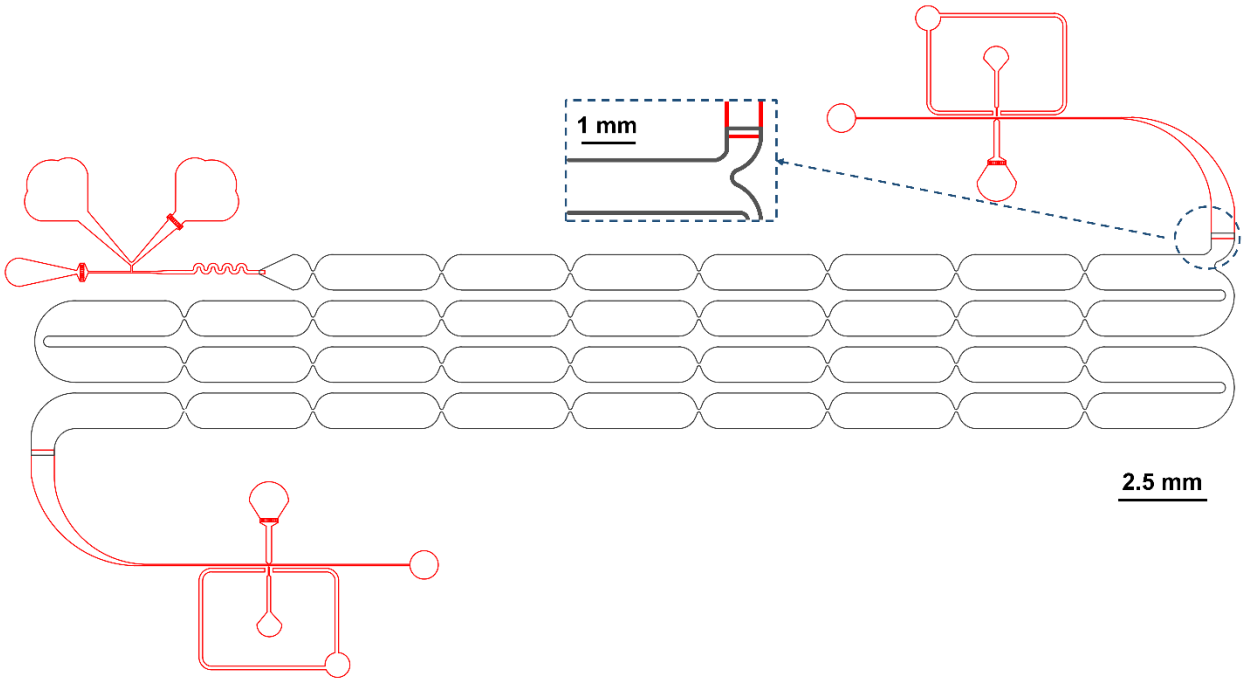

Figure A-2 A representative of non-selected device design for MBA/MNase-seq sample preparation in Chapter 4. The shown design was named "4R", as it contained 4 rows of thick, 160-µm delay channels in total.

Appendix B

Biological Related Information and Results

Table B-1 Sequences of primer pairs for the positive and reference loci used in the MBA/MNase device development in Chapter 4.

Primer name	Sequences (5’ to 3’)
CD47 enhancer forward primer	TGATGACACCGGTGATATGGT
CD47 enhancer reverse primer	GGCATTTTATCTTCCACTCTCCC
DZIP3 forward primer	CTCTGCATAACCACTAGGTGGCA
DZIP3 reverse primer	TTTTTAAACAAGTGCAGTCGTGTGG
C19 intergenic forward primer	AGCTTGTCTTTCCCAAGTTTACTC
C19 intergenic reverse primer	TAGCTGTCGCACTTCAGAGGA

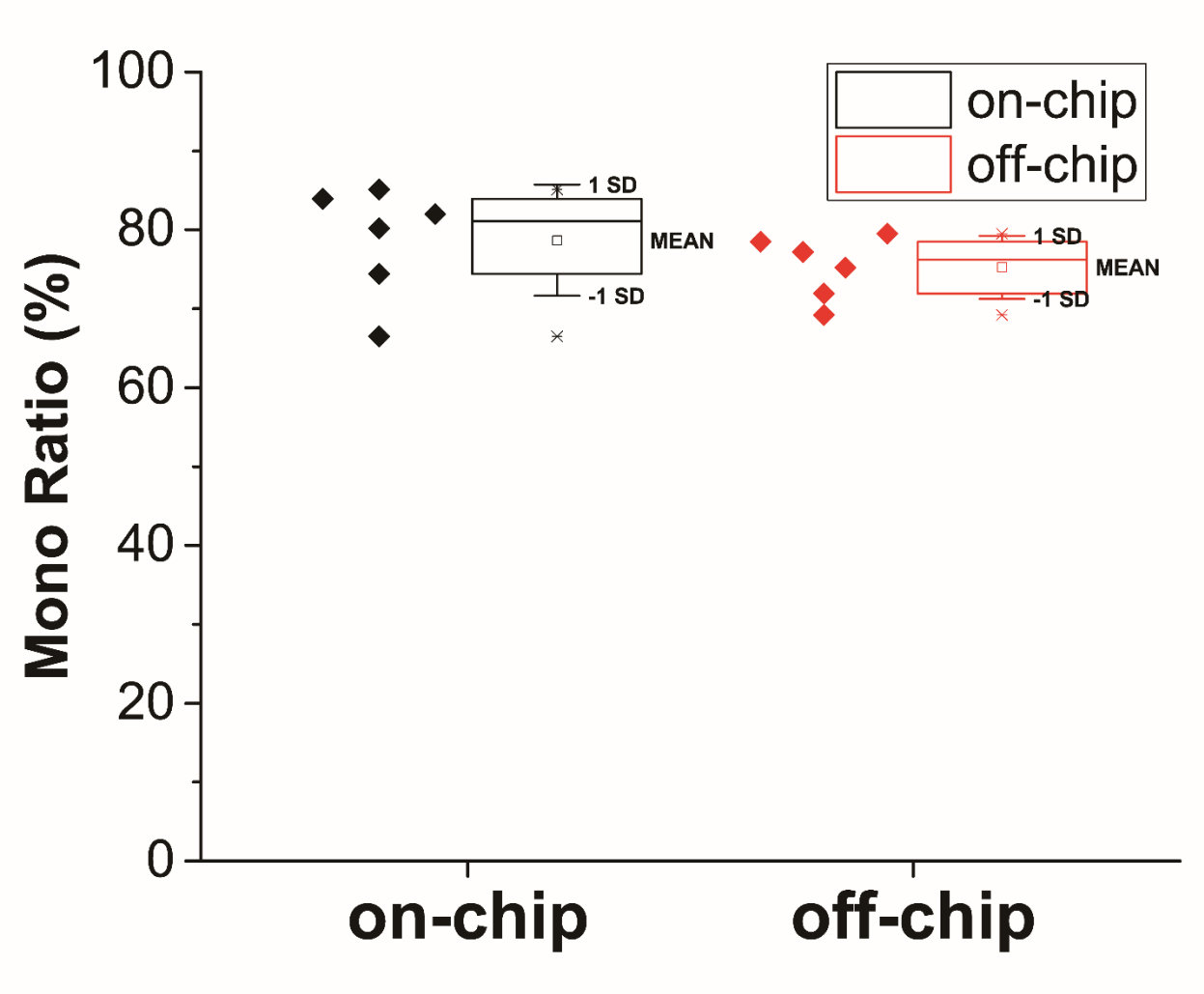

Figure B-1 Mononucleosome yields from all Jurkat samples processed for sequencing with the data points illustrated on the left and corresponding box plots on the right for both on-chip and off-chip samples. Each sample was 50-µL suspension containing 500,000 cells. The whiskers of the box plots represented one standard deviation. The top and bottom box limits represented third and first quantiles of the corresponding data. The lines in the middle of the boxes represented means of the data. The small squares in the middle of the boxes represented medians of the data.

Figure B-2 Digestion pattern of K-562 cells MBA DNA of one replicate. Notice only a small peak < 250 bp existed in MBA DNA while majority was at higher base pair approaching to the higher internal standard, in this case the "mono ratio" was 9.2%. Level of digestion similar to this was expected to be qualified for the limited digestion to preserve linker DNA or nucleosome-free accessible regions. The two high peaks were internal standards at 35 bp and 10380 bp, respectively. The green and blue dashes were location cursors of the Bioanalyzer software, which had no effect on the profile.

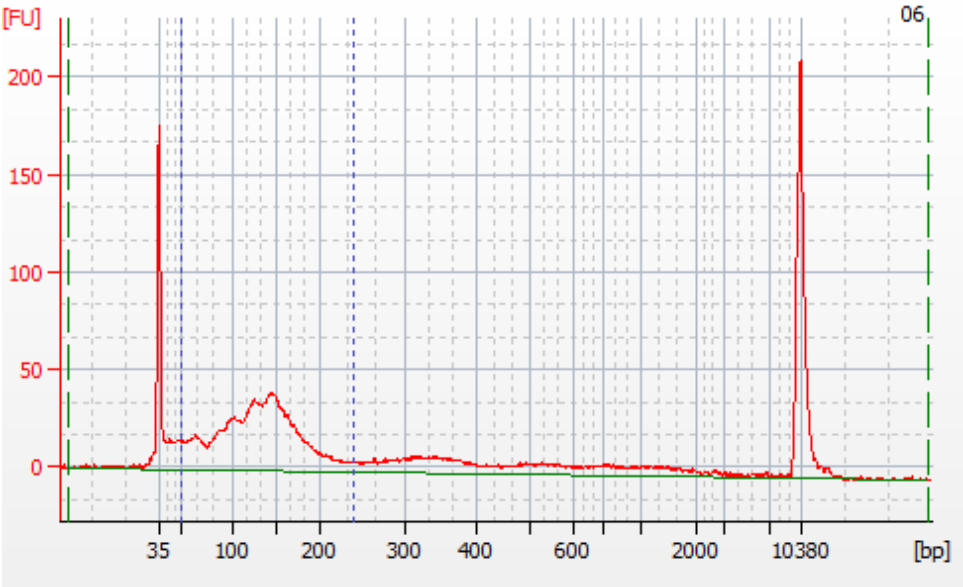

Figure B-3 Digestion pattern of K-562 cells MNase DNA of one replicate. Notice that majority of the DNA was digested to mononucleosomal length with only small di-nucleosome peak. Level of digestion similar to or higher than this was expected to be qualified for MNase-seq. The two high peaks were internal standards at 35 bp and 10380 bp, respectively. The green and blue dashes were location cursors of the Bioanalyzer software, which had no effect on the profile.

Figure B-4 Digestion pattern of buffy coat cells MBA DNA of one replicate. Notice only a small peak < 250 bp existed in MBA DNA while majority was at higher base pair approaching to the higher internal standard, in this case the "mono ratio" was 21%. Level of digestion similar to this was expected to be qualified for the limited digestion to preserve linker DNA or nucleosome-free accessible regions. The two high peaks were internal standards at 35 bp and 10380 bp, respectively. The green and blue dashes were location cursors of the Bioanalyzer software, which had no effect on the profile.

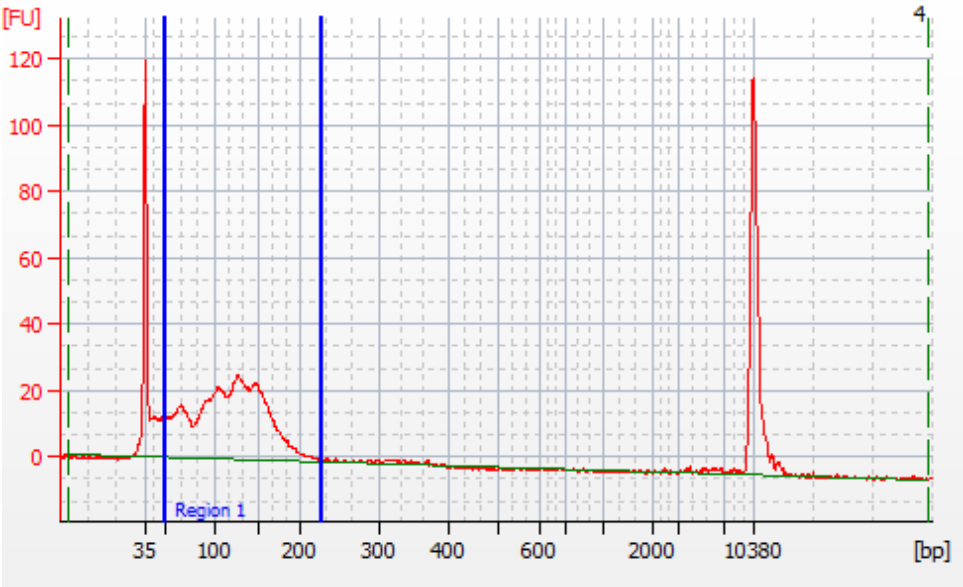

Figure B-5 Digestion pattern of buffy coat cells MNase DNA of one replicate. Notice that majority of the DNA was digested to mononucleosomal length with only small di-nucleosome peak. Level of digestion similar to or higher than this was expected to be qualified for MNase-seq. The two high peaks were internal standards at 35 bp and 10380 bp, respectively. The green and blue dashes were location cursors of the Bioanalyzer software, which had no effect on the profile.

Table B-2 Sequences of primer pairs for the targeted and reference loci used in the development of droplet microfluidic ChIP in Chapter 5.

Primer name	Sequences (5’ to 3’)
Brg1_TSS forward primer	TTGGCGAAGCTGCGATCGGG
Brg1_TSS reverse primer	AGGGGACCGCTAATGCCCGT
MYT1_TSS forward primer	CCTGCCGTGTGCTGTTTTT
MYT1_TSS reverse primer	CACAACATGTCCCCTGGAATC
SAT-alpha forward primer	AAGGTCAATGGCAGAAAAGAA
SAT-alpha reverse primer	CAACGAAGGCCACAAGATGTC
C19 intergenic forward primer	AGCTTGTCTTTCCCAAGTTTACTC
C19 intergenic reverse primer	TAGCTGTCGCACTTCAGAGGA

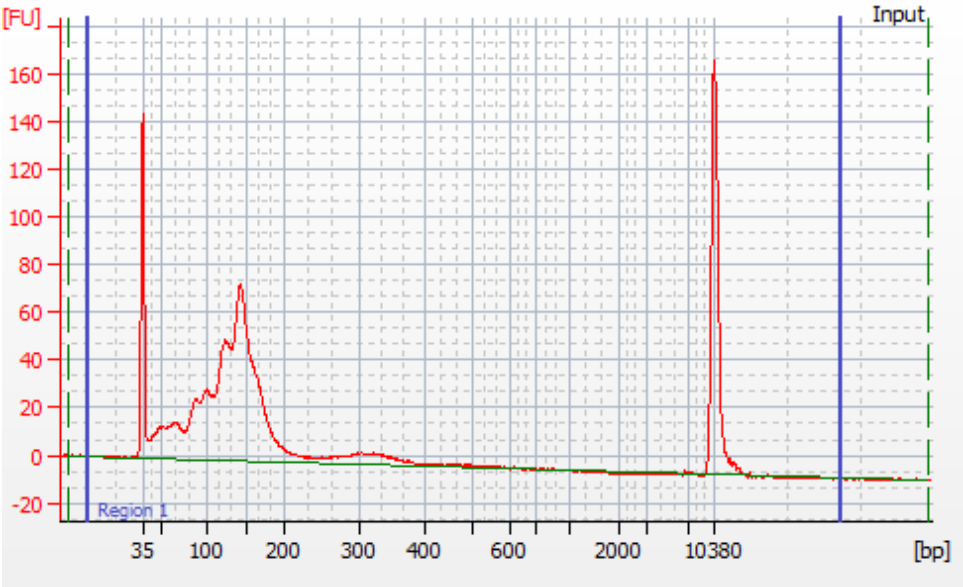

Figure B-6 Digestion pattern of the input sample used in ChIP in tube mimicking on-chip conditions targeting H3K4me3. Notice that majority of the DNA was digested to mononucleosomal length with only small di-nucleosome peak. Mono peak consisted of 80.6% of the entire digestion profile. The two high peaks were internal standards at 35 bp and 10380 bp, respectively. The green and blue dashes were location cursors of the Bioanalyzer software, which had no effect on the profile.

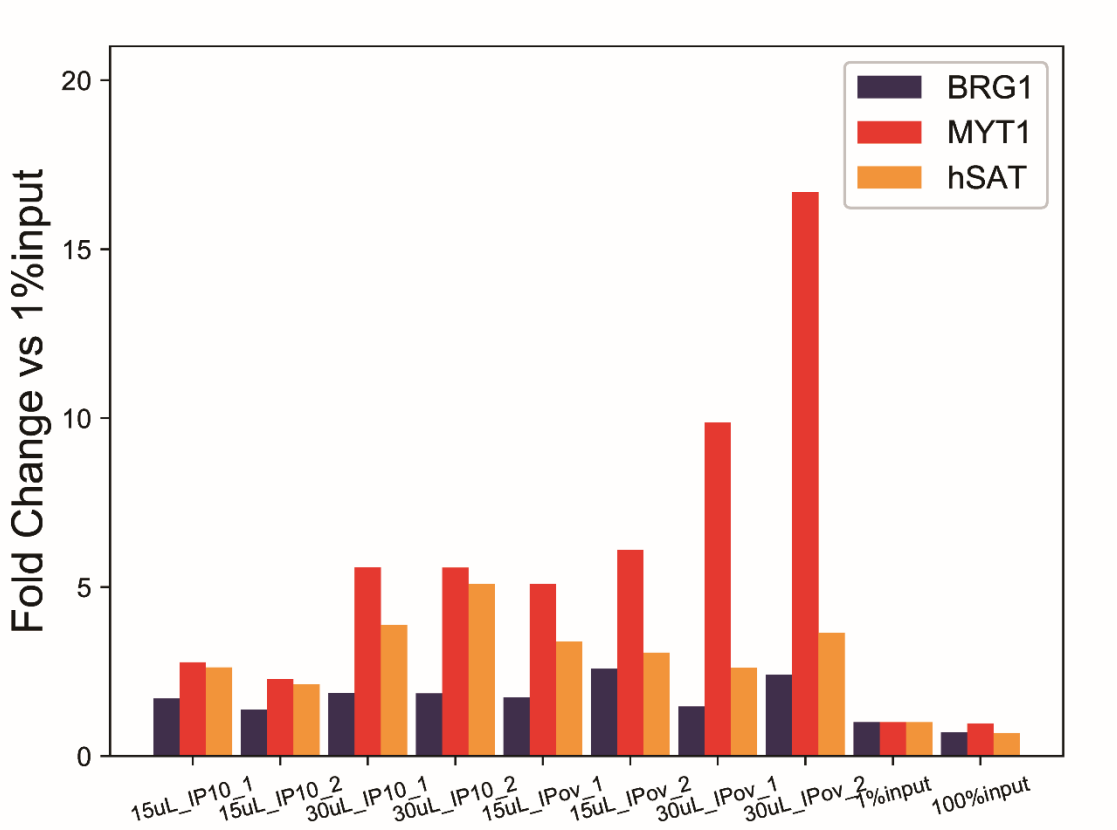

Figure B-7 Fold change results of ChIP using 2-µg anti-H3K27me3 antibody per reaction. 15uL_IP10_1 stood for ChIP with 15-µL Dynabead slurry, incubating 10 minutes for immunoprecipitation, replicate 1. "ov" stood for ChIP overnight. The 1% input came from preparing enough aliquots of fragmented chromatin for each ChIP reaction together and took 1% out of one aliquot of chromatin from that. BRG1 stood for fold change of *Brg1_TSS* normalized to *C19_intergenic* in ChIP DNA compared to the input sample F3_1_1%inp, MYT1 stood for fold change of *Myt1_TSS* normalized to *C19_intergenic*, hSAT stood for fold change of *SAT-alpha* normalized to C10_intergenic. Notice the highest fold change of *Myt1_TSS* and similar fold change of the other two non-specific loci obtained in the condition of 30-µL Dynabead slurry incubating overnight. This condition was the one described as "ChIP in tube mimicking on-chip conditions". Fold changes of three gene loci all close to 1 in 100% input showed that the chromatin fragmentation was consistent across replicates and parts of replicates, showing that simultaneously lysed and digested native chromatin suspension was homogenized well.

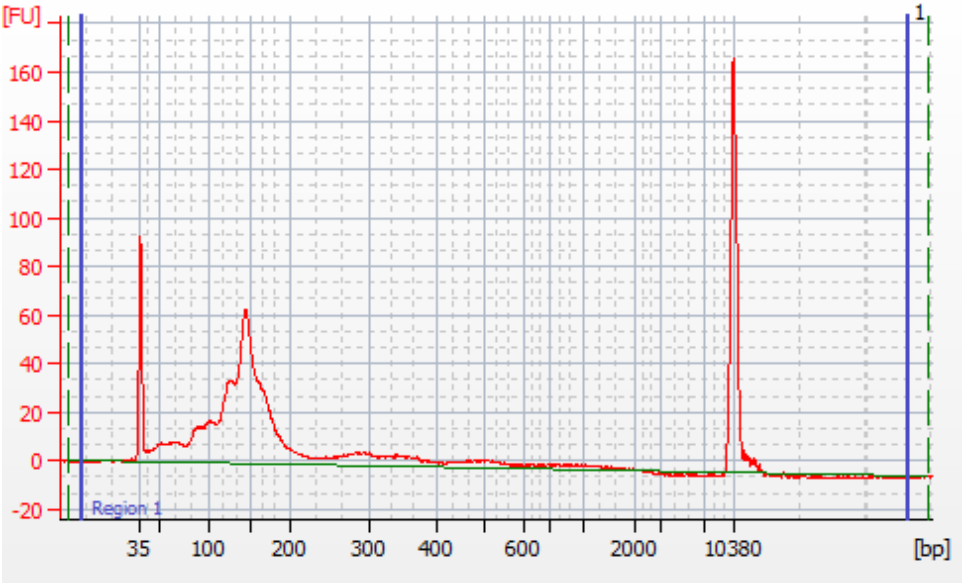

Figure B-8 Digestion pattern of the F3_1_1%inp input sample used in ChIP of three formats. Notice that majority of the DNA was digested to mononucleosomal length with only small di-nucleosome peak. Mono peak counting from 75 bp to 235 bp consisted of 74.1% of the entire digestion profile. The two high peaks were internal standards at 35 bp and 10380 bp, respectively. The green and blue dashes were location cursors of the Bioanalyzer software, which had no effect on the profile.

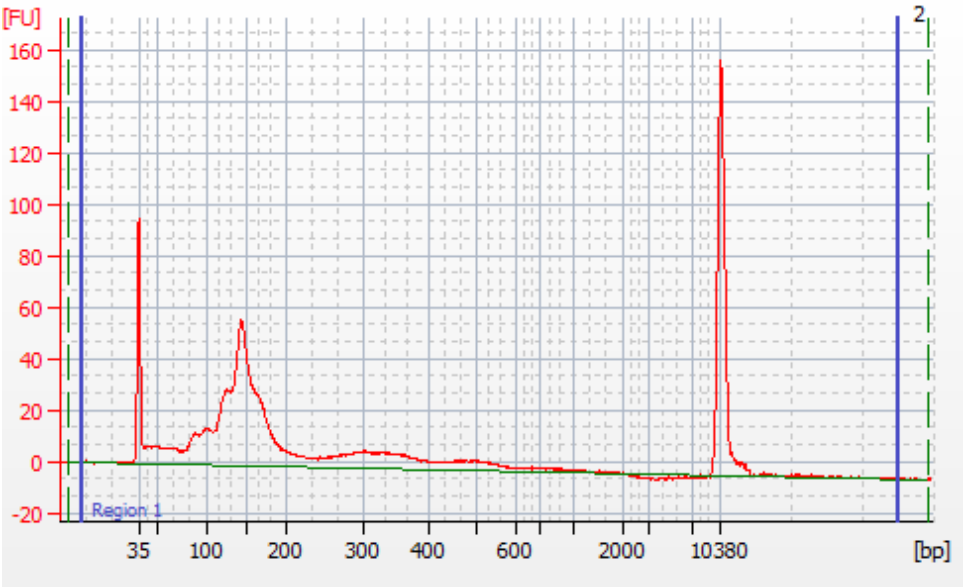

Figure B-9 Digestion pattern of the F3_2_1%inp input sample used in ChIP of three formats. Notice that majority of the DNA was digested to mononucleosomal length with only small di-nucleosome peak. Mono peak counting from 75 bp to 235 bp consisted of 71.8% of the entire digestion profile. The two high peaks were internal standards at 35 bp and 10380 bp, respectively. The green and blue dashes were location cursors of the Bioanalyzer software, which had no effect on the profile.

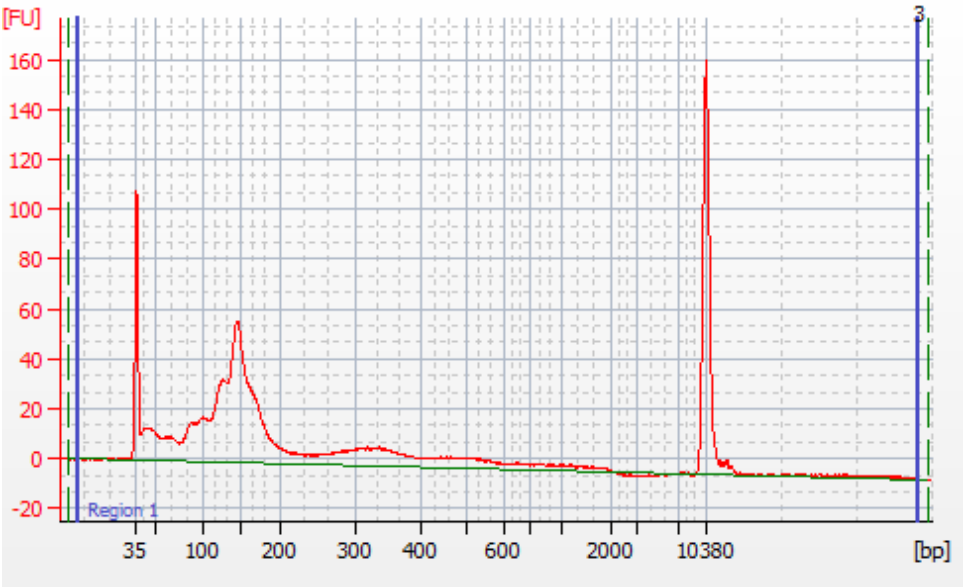

Figure B-10 Digestion pattern of the In_100%inp input sample used in ChIP of three formats. Notice that majority of the DNA was digested to mononucleosomal length with only small di-nucleosome peak. Mono peak counting from 75 bp to 235 bp consisted of 66.2% of the entire digestion profile. The two high peaks were internal standards at 35 bp and 10380 bp, respectively. The green and blue dashes were location cursors of the Bioanalyzer software, which had no effect on the profile.