DEVELOPMENTS TOWARDS SINGLE MOLECULE AND SINGLE CELL GENETIC AND EPIGENETIC ANALYSIS TECHNOLOGIES

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Harvey Chen Tian

May 2017

DEVELOPMENTS TOWARDS SINGLE MOLECULE AND SINGLE CELL GENETIC AND EPIGENETIC ANALYSIS TECHNOLOGIES

Harvey Chen Tian, Ph.D.

Cornell University 2017

New discoveries in the realm of biology often come hand-in-hand with innovations in technology. Over the last two decades, advancements in microfabrication and microscopy have opened biological phenomena to be studied in fundamental units rather than as a collective entity. From these efforts, the axioms upon which we built our fundamental concepts in anatomy and physiology have been reshaped. We now understand that even within a subpopulation of cells taken from the same tissue, cell-to-cell variations exist in both phenotypic and genotypic expression. These variations are found to play critical roles in disease state and progression, thus, research into single cells and single molecules are a necessity to improving treatment.

In this work, methods of single molecule analysis of native human chromatin fibers are presented as well as methods and devices for multi- and single cell genome analysis. As the handling of single molecules and single cells

are facilitated by microfabricated devices, discussed herein are their designs, fabrication methods, operational procedures, working principles, and experimental results representative of their biologically relevant impact. Our findings point towards potential platform technologies in high-throughput chromatin linearization for fluorescence based epigenetic mapping and single cell whole genome amplification with reduced amplification bias and improved genome coverage.

BIOGRAPHICAL SKETCH

Harvey Chen Tian was born to Dr. Zheng Rong (Ryan) Tian and Hua (Tina) Zhong in Shanghai, China on 6 July 1988. At the age of 18 months, Harvey traveled with his mother to the United States of America to meet with his father who had just begun his Ph.D. studies at the University of New Orleans, Louisiana. Eventually, Harvey's father transferred to the University of Connecticut in Storrs, Connecticut where Harvey and his mother followed. Here, Harvey's sister, Linda, was born on 12 December 1995 who would later become Harvey's closest friend in coming years. As Harvey's father's career began to take hold, Harvey and his family would move five more times, living in California, Albuquerque, and New Mexico before finally settling down in Arkansas.

Harvey completed his high school education in the prestigious magnet and boarding high school known as the Arkansas School for Mathematics, Sciences, and the Arts located in Hot Springs, Arkansas. Here, his interest for biology and engineering led him to pursue a bachelor's degree in biomedical engineering upon being accepted to attend the Georgia Institute of Technology. While studying as an undergraduate student at Georgia Tech, Harvey worked in the laboratory of Professor Rudolph Gleason of the biomedical engineering department. Although Harvey gained a substantial amount of laboratory and

research experience from his work in Dr. Gleason's lab, he realized that his passions in research were in developing technology for biomedical applications rather than studying biological phenomena. After applying to Cornell as a Master of Engineering student, Harvey reached out to nanobiotechnology groups on campus to which Professor Harold Craighead responded. Professor Craighead welcomed Harvey to his lab and gave him the opportunity to work with outstandingly talented post-doctoral fellows in his lab, who would directly mentor Harvey throughout his time as a Master student. Harvey quickly assimilated himself into Harold Craighead's lab and contributed towards projects that would set him on the path of his Ph.D.

Finally, with the help of Harold Craighead and with the work he accomplished during his time as a Master student, Harvey applied for and was accepted into the biomedical engineering Ph.D. program on 16 April 2012. Harvey continued his research under the mentorship of Harold Craighead, calling Cornell and Ithaca his home for the next five years.

Harvey hopes to carry the traditions, values, and kindness that he was shown at Cornell with him throughout his life as he aspires to give back to the people and community that bestowed so much upon him.

DEDICATION

To my deceased grandfather, whose legacy as a scholar led my family onto the path of knowledge and education. And to my parents, for their persistent sacrifice, support, and love.

ACKNOWLEDGEMENTS

My aspiration of studying at Cornell was preceded by my late grandfather, a true scholar whom always dreamed of attending this great University. Although he was unable to fulfill his dreams due to the political climate of his time, I am honored to know that I have completed this piece of his legacy in his place.

For this entire opportunity, I sincerely thank my advisor, Harold Craighead. Thank you Harold for believing in me before anyone else at Cornell and for continuing to have faith in me even through times which I doubted myself. Your patience, encouragement, and mentorship guided me through my journey as a graduate student. You not only showed me kindness and respect, you were willing to listen to my concerns and opinions. Through you, I learned to become a better communicator, scientist, and thinker. You challenged me to grow as a person and for that I am grateful.

To my fellow lab mates Juraj Topolancik, Aline Cerf, Jaime Benitez, and Kylan Szeto, thank you for helping me along my journey, and more importantly, thank you for being my closest friends both inside and outside of lab. Your encouragement and advice formed the foundation of my development through

graduate school. I greatly miss working with you all and I hope that I will have the pleasure of working with individuals of your caliber throughout my career.

To my friends in Ithaca, thank you Julie Kohn and Mary Clare McCorry for your exceptional friendship, companionship, and comradery throughout my PhD. You are my dearest friends with whom I started the Ph.D. program and with whom I share many cherished memories.

To my mother, father, and my sister Linda, thank you for being my strongest support system. Your words of wisdom, encouragement, and many, many hours over the phone have been the cornerstone to my successful completion of this journey. Your persistent love is immeasurable and I cannot ever thank you enough for everything you have given me.

To my beautiful girlfriend, Aurea Stephany Rivera-Burgos. Your love, patience, encouragement, humor, and wisdom has helped me through the most stressful times in my Ph.D. Thank you for the joy you have brought to my life and for making me feel at home away from home.

To the Adirondack Mountains, thank you for your trails and breathtaking views. Solo-hiking your peaks has been the single most healing and meditative practice I have experienced.

Finally, thank you to the National Institutes of Health of the United States of America for funding my research, without which, this work would not have been possible.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

OVERVIEW

DNA, widely regarded as the fundamental building block of life, has been extensively studied as both a physical molecule as well as an information carrier for nearly a century. Only recently have advances in micro- and nanotechnology allowed us to precisely characterize and probe the information contained within DNA. These advancements have led to the understanding that not only are heritable chemical modification of the genome important to gene regulation and gene expression, but that each individual cell can differ from one another within a tissue type. However, studying these phenomena requires new tools to be developed. To this end, the microfabricated technologies we have developed to improve upon existing methods are presented.

Chapter 2 covers the motivations behind developing methods for chromatin linearization and imaging by describing the basic principles of epigenetics and existing methods of epigenetic analysis. We then discuss recent advances in epigenetic analysis and chromatin linearization before moving into Chapter 3, where we describe in detail our own technique producing ordered arrays of native chromatin using molecular combing and transfer printing. Together Chapters 2 and 3 mark part I of this dissertation.

Part II begins with Chapter 4, where microfluidic based cell processing devices are analyzed for their benefit unique advantages in handling few and single cells. We present our device, a simple, valveless, micropillar array based device cast from PDMS designed for cell capture, lysis, and DNA extraction. In this chapter, we characterize various parameters of the device and assess for DNA capture efficiency of multiple and single cells.

This work then segues into Chapter 5, where we examine single cell technologies in more detail as compared to our device. By redesigning the micropillar array device for single cell capture, we utilize the device as a platform for on-chip whole genome amplification.

Finally, Chapter 6 is where we discuss future directions for the microfluidic device and potential experiments that will demonstrate the unique advantages conferred by our chip as a DNA amplification platform device as compared to existing single cell technologies.

CHAPTER 2

EPIGNETICS AND DNA LINEARIZATION

2.1 INTRODUCTION TO CHROMATIN AND EPIGENETICS

With roughly 3 billion bases comprising the human genome, the DNA from a single cell would span nearly 2 meters in length if stretched out from end to end.[1] In order for all of this DNA to be packaged within the cell nucleus, DNA exists in eukaryotic cells as a part of chromatin. Chromatin refers to the higher order structures of DNA wrapped around histone protein complexes. Comprised of an octamer formed by histone proteins H2A, H2B, H3, and H4, the histone core complex is bound together by 146 bases of DNA with the help of linker protein histone H1.[2] Together, this structure forms what is known as a nucleosome, the fundamental repeating unit by which chromatin is built. Between each nucleosome are segments of linker DNA that range in length from tens of bases up to several kilobases.[3]

However, rather than remaining fixed in a permanent configuration, chromatin is a dynamic molecule constantly undergoing conformational changes as cellular processes such as gene transcription and cellular division occurs.

Known as chromatin remodeling, the chromatin hierarchal structure acts as a controller mechanism for gene regulation by physically modulating accessibility of the genome to transcription related proteins such as RNA polymerase and transcription factors.[4] If a gene promotor region is inaccessible due to the structure of the local chromatin, RNA polymerase is unable to initiate transcription and form an open complex. In such cases, messenger RNA (mRNA) for a particular gene would be unable to be transcribed and subsequently translated into a given protein even though the gene for such a protein exists within the genome. This is the case for heterochromatin, a tightly wound configuration of chromatin that consists of inactivate gene sites and telomeres packaged into tertiary order 30nm diameter fibers with the help of histone H1 and is heavily localized to the walls of the cell nucleus. The existence of heterochromatin has been suggested to serve roles of chromatin preservation as well as epigenetic inheritance.[5-6] In contrast to heterochromatin, euchromatin is a much more accessible conformation of chromatin containing actively transcribed genes and nucleosomes loosely configured in a "beads-on-a-string" format. Euchromatin represents the default configuration for the vast majority of the human genome and unlike heterochromatin, euchromatin is active in gene transcription.[7]

The mechanisms by which chromatin remodeling form heterochromatin versus euchromatin are controlled through a set of covalent modifications. These modifications can exist in the form of DNA methylation, hydroxymethylation, and oxidation or histone modifications in the form of acetylation, methylation, phosphorylation, and ubiquitination. Collectively, these chemical modifications are referred to as the epigenetic state or the epigenome of a cell. Research into epigenetics have revealed the critical relation between epigenomics and disease. In humans, numerous conditions including heart disease, autism, obesity, and various forms of cancer have been found to be correlated to epigenetic states and modifications.[8-10]

A prominent example of one such type of correlation is with CpG island hypermethylation. Denoting a repetitive sequence of cytosine bases and guanine bases connected via the DNA phosphate backbone, CpG sites can cluster into high density within the genome regions known as CpG islands that range in size from hundreds of bases up to a few kilobases. The importance of CpG islands is that in humans, 60% of gene promotor regions contain CpG islands.[11] As such, hypermethylation of cytosines within CpG islands can lead to gene silencing.[12-13] For example, when genes such as MLH1 or BRCA1 are suppressed via this process, cancer onset can occur resulting in colon or breast cancers respectively.[14-15] Furthermore, suppressing

transcription via DNA hypermethylation can lead to hosts of other diseases including those of neurological and autoimmune pathogenesis.[8]

Consequently, an active area of medicine aims to improve our understanding of the epigenome and its ability to affect health and disease states. Efforts in this field contribute towards mapping the human epigenome and the complex relations between epigenetic modifications and disease. In the future, with the epigenome mapped and a fuller understanding of the ties between various epigenetic modifications and diseases, the opportunity to create diagnostics and eventually even therapeutics based on the epigenome may come to light.

2.2 PRIMER ON EPIGENETIC MAPPING

Currently, the standard in epigenetic mapping utilizes two techniques in conjunction with DNA sequencing. The first is bisulfite sequencing, which is used to map DNA methylation positions on the genome, and the second is chromatin immunoprecipitation (ChIP) which is used to map the location of specific histone modifications on the genome. The cornerstone of bisulfite sequencing based epigenetic mapping techniques rely on converting unmethylated cytosine bases into uracils through sulfonating the cytosine

molecule, hydrolytic deamination, and finally desulfonation. This process, however, is disrupted in the presence of 5′ methylation. In such cases, a CH_3 methyl group binds to the 5-prime position carbon atom on the cytosine molecule preventing the sulfonation of the molecule. Thus, sodium bisulfite converts all non-methylated cytosine bases into uracils while methylated cytosines remain. Through subsequent DNA amplification and sequencing, the sequence of bisulfite converted DNA can be compared to the sequence of identical genome regions that have not been bisulfite converted to determine the genomic positions of cytosine methylation.

On the other hand, in ChIP based sequencing (ChIP-seq), native chromatin is chemically treated with formaldehyde to crosslink histone complexes to the local DNA. The DNA is then fragmented either physically as with sonication or enzymatically. Histone modifications of interest are then selected from the suspension of nucleosomes by binding epigenetic modifications of interest to substrates or particles functionalized with target-specific antibodies. The arrested nucleosomes containing histone modifications of interest are then isolated from solution and purified. This process is referred to as immunoprecipitation. The chromatin is eluted from the beads via heat denaturation and the DNA that is bound to the nucleosomes are then unlinked from the histone proteins with the help of proteinases such as proteinase K. The

released DNA represents the genome location of the histone modifications selected by immunoprecipitation and can be sequenced to map the epigenetic mark back onto the reference genome.

Although both of epigenome mapping approaches offer unique advantages, the utilities of each are far from sufficient for use in clinical diagnostics. One such shortcoming of these two techniques is that although they offer a high degree of accuracy, they are considerably time consuming and require on the order of millions of cells.[16] Despite recent advances that allowed both ChIP-seq and bisulfite sequencing to reliably operate from only a few thousand cells, the sensitivity limits of both techniques still bar them from being used in low-cell count scenarios.[17] For example, cancer metastasis is initiated by aggressively mobile cancer cells within the body known as circulating tumor cells (CTC). These rare cells often exist at ultralow concentrations of less than 10 cells per milliliter of whole blood.[18] Hence, determining the epigenetic behavior of patient CTCs remains highly challenging not only from the perspective of capturing the CTCs but also in terms of analysis methods with enough sensitivity.

Furthermore, there are fundamental limitations to ChIP-seq and bisulfite sequencing that cannot be removed by optimization of current protocols. ChIP-

seq is limited in only being able to serially interrogate one subset of histone modifications at a time. The information from nucleosomes containing multiple sets of epigenetic modifications would be lost in immunoprecipitation. Similarly, bisulfite sequencing can offer high resolution methylation mapping, however, the dynamic behavior of the epigenetic state is lost in both analysis methods as they require permanent alteration of the template genomes such as crosslinking and fragmentation in immunoprecipitation and DNA degradation in sodium bisulfite conversion. In this vein, exploring alternative epigenetic mapping techniques are critical to the overall development of epigenomics-based diagnostics. One such alternative to ChIP-seq and bisulfite sequencing is to probe for epigenetic marks of interest directly on the genome with the aid of technologies from the field of nanobiotechnology.[16,19]

2.3 SURVEY OF DNA & CHROMATIN LINEARIZATION TECHNIQUES

Itself an eclectic field derived from the broader field of biophysics, nanobiotechnology has made significant impacts in the biological research since its inception half a century ago. Over the last three decades, advancements in micro- and nanofabrication have allowed scientists to begin analyzing biological phenomena at its most fundamental level: single molecules. The

power of using high-resolution single molecule techniques to directly interrogate epigenetic marks on the genome without needing to destroy or alter the genome confers several advantages over methods such as ChIP and bisulfite sequencing, which can only indirectly infer the epigenetic state at hand. In this section, several established methods for single molecule handling of DNA and chromatin are evaluated for their individual strengths and weaknesses in application towards epigenetic mapping.

2.3.1 Magnetic and Optical Tweezers

Using either magnetic forces or optical traps to micromanipulate beads tethered to DNA molecules, magnetic and optical "tweezers" have been used to study the mechanical properties of DNA as early as 1992 and 1996 respectively by Smith S. B. et al.[20-21] Briefly, magnetic tweezers use magnetized microparticles that are surface functionalized with proteins such as streptavidin, which forms a strong bond to biotin. Biotin can then be hybridized the end of a chromatin or DNA molecule. Either one or both ends of the biological molecule can be tethered to a bead or a planar surface and the bead can be manipulated with the magnetic field created by a set of permanent magnets. Repositioning the magnets allows for rotational and directional

control over the magnetic microparticle which translates into forces acted upon the biological molecule. In similar fashion, optical tweezers also rely on a protein-protein interaction to anchor the biological molecule to a microsphere. However, rather than relying on a magnetic field, a laser is focused through a high numerical aperture objective lens creating a narrow beam waist that is able to impart the momentum of scattered incident photons onto a dielectric microparticle.

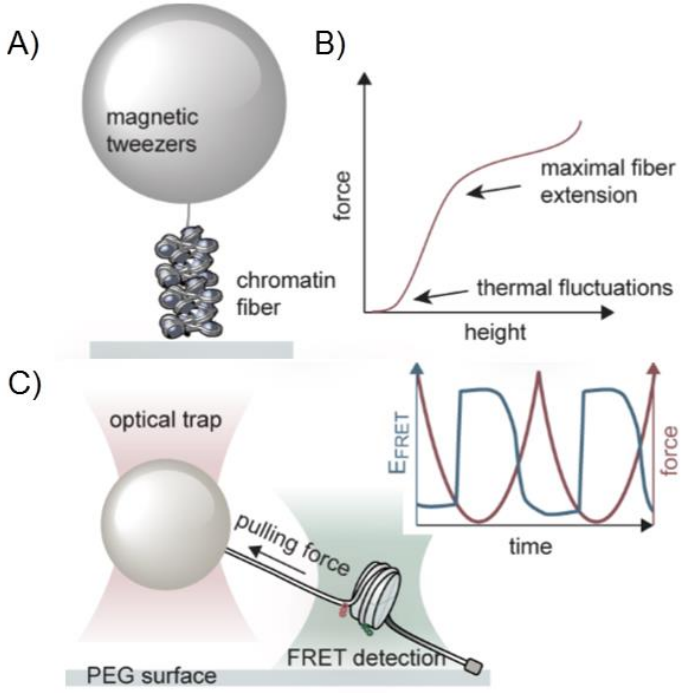

Figure 2.1 – Illustration of Force Measurement Experiments Performed with Magnetic and Optical Tweezers

(Modified from Fierz B., 2016)

(A) Magnetic microsphere coated with streptavidin can be tethered to a chromatin with a biotinylated end. By pulling the microsphere directionally along the z-axis, (B) chromatin fiber thickness can be measured as a function of the length of DNA linkers between nucleosomes until all nucleosomes have been pulled off. (C) The principles of Fluorescence Resonance Energy Transfer (FRET) can be used to monitor separation between the two points to denote nucleosome unraveling. As force is increased, the distance between two FRET molecules grows thereby lowering the fluorescence intensity of the pairing.

Both optical and magnetic tweezers offer the advantage of being able to elongate DNA or chromatin with a high degree of control in the force applied to the molecule.[22] With sub-pN force resolution, past research has used tweezer-based methods to determine contour length of DNA, measure tensile forces of single nucleosomes, measure forces required to unzip double stranded DNA, alter the conformational state of single chromatin molecules through elongation, and make real-time observations of protein dynamics.[23-24] Examples of the workflow of a few such studies are illustrated in ***figure 2.1*** above.[24] Unfortunately, tweezer based methods are at a disadvantage when conjugating microspheres to unknown sequences such as native chromatin from cells. To date, the most apt demonstration of using optical tweezers to

elongate native chromatin was done by Oana H. and colleagues in 2014 as shown in ***figure* 2.2** below.[25] In this study, authors relied on functionalizing histones to tether the chromosome to the dielectric microparticle, which precluded the ability to image chromosome molecules without folding over of the chromosome at the tethered end.

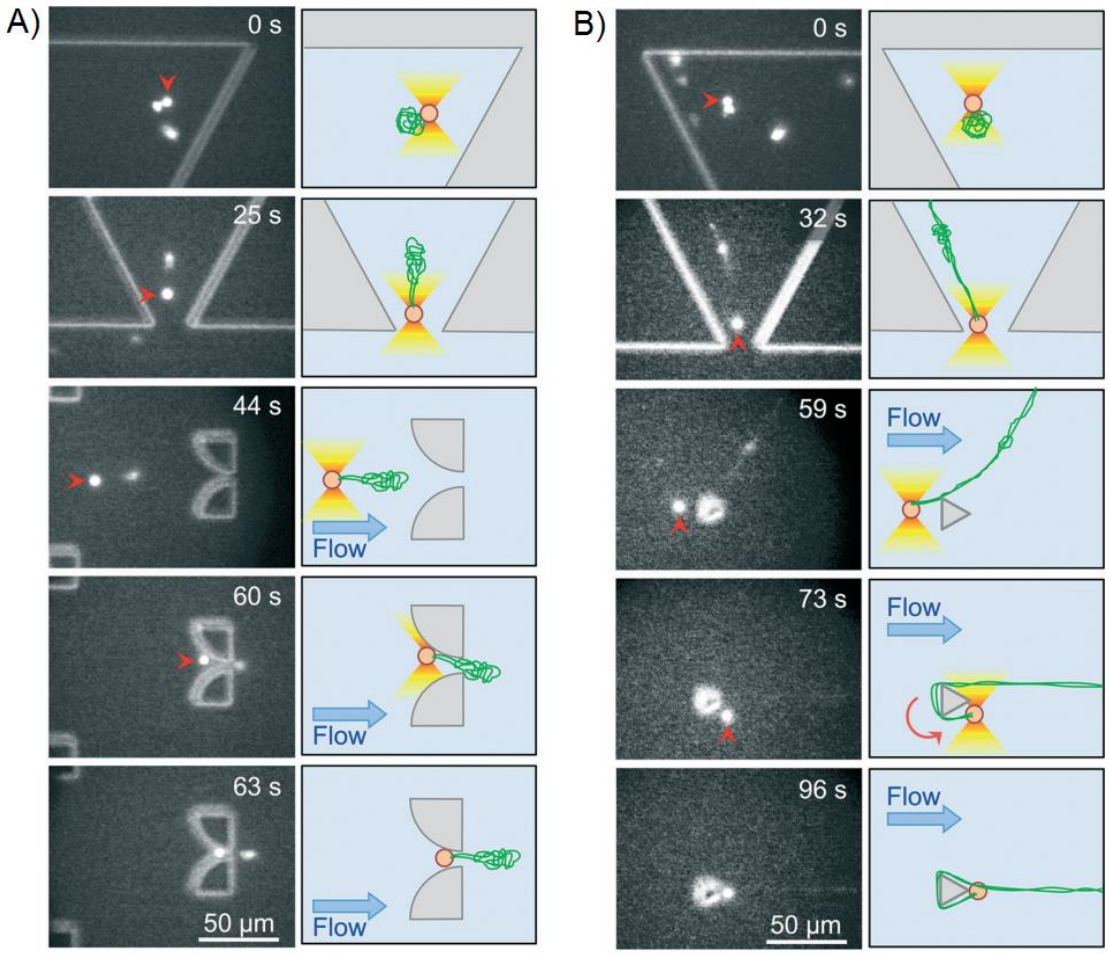

Figure 2.2 – Optical Tweezer Based Elongation of Single Chromatin Fibers via Microsphere Anchoring to PDMS Posts

(Adapted from Oana H. et. al., 2014)

Cells are lysed within a triangular compartment in microfluidic device. Functionalized microspheres that bind non-sequence specifically are used to fish out and maneuver the chromatin. (A) Long chromatin fibers are easily entangled during translocation and become difficult to pass through twin polydimethylsiloxane (PDMS) posts designed to anchor the microsphere. (B) An alternative design where less entangled chromatin is looped around a single pillar due to the difficulty in directing a long fiber through gap between two close-set microposts.

All-in-all, although tweezer-based studies have been substantial to the overall understanding of DNA structure, nucleosome mechanical properties, and DNA-protein dynamics, optical and magnetic tweezers face many challenges including ultra-low throughput and limited molecule lengths when considered as a tool for epigenetic mapping.

2.3.2 Nanopore Based Technologies

The essence of nanopore based technologies can be described as using voltage differentials to translocate DNA or chromatin through a hole on a thin dielectric

membrane and measuring changes in the electrical signal of the substrate to determine the local structure of the molecule as it passes through. Since different nucleic acid bases translocate at different speeds and impede the ionic-current to differing degrees, nanopore technologies are used as a method of sequencing of DNA and RNA. Although it widely accepted that pore sizes should prevent looping events from traversing the pore during sequencing, there has yet to be an established standard optimal pore size.[26] As the spatial resolution of nanopore technologies depend on both the thickness of the membrane and size of the pore, researchers have made efforts to reduce membrane thicknesses of various solid-state nanopores down to sub-nanometer thick silicon nitride and even single-atom layer graphene membranes.[26-28] Transmembrane proteins such as α-hemolysin, MspA, and Phi29 are commonly used to create biological pores and in recent years, they have been incorporated onto solid-state membranes to form hybrid nanopore devices. The advantage of using biological pores is improved signal-to-noise ratio over solid-state pores that can experience high noise levels. This improved signal can be further facilitated by the ratcheting function of the proteins such as Phi29 to modulate the speed of DNA translocation through the pore.[29-30]

Even more recently, DNA methylation has been observed with the use of nanopore based technologies. In a study done by Shim. J. et. al. in 2013, methyl-

binding domain protein-1 (MBD1) was used to label methylation sites along a strand of methylated DNA.[31] The translocation current of the MBD1-labeled DNA was than measured as it was passed through a 12 nm diameter solid-state pore in 20nm thick silicon nitride. The results of this experiment showed the ability to differentiate unmethylated DNA from their control samples versus methylated DNA with high fidelity. However, the authors are unable to recount the exact number of methylation sites as translocation speed still proves to be an issue regarding reading spatial resolution of the DNA molecule. Molecular dynamic simulations have proposed functionalizing the inner pore edges with DNA bases to improve sequencing accuracy and slow DNA translocation times, but these theories have yet to see any validity experimentally.[28] Furthermore, removing DNA interactions with the membrane wall at the pore interface still presents a challenge.[28,32,35] Localized non-covalent blocking has been proposed, but has yet to see widespread use.[33,36] Finally, although proof of concept experiments such as those from Soni G. V. and colleagues have shown the ability to differentiate between mononucleosomes versus dinucleosomes using 20nm diameter pores in silicon nitride, longer chromatin molecules have not been demonstrated to be passed through nanopores for identifying or mapping epigenetic marks.[34] Consequently, although nanopore technologies represent a time-efficient means of analyzing methylation marks, it in DNA, much work has yet to be done before nanopores can be used as a basis for

epigenetic mapping. The possibility of simultaneous methylation detection and DNA sequencing has not been evidenced and longer chromatin reads are not yet possible.

2.3.3 Nanochannel and Confinement Technologies

Through using physical barriers to spatially confine DNA and chromatin molecules into linear orientations, nanochannel and nanoconfinement based technologies have emerged as feasible approach to directly interrogating epigenetic marks of interest. Generally, this involves labeling DNA or chromatin with fluorescent tags specific to epigenetic marks of interest and using high resolution imaging to determine the position of an epigenetic event on the molecule. This positional information relative to the DNA or chromatin molecule can then be transposed back the genome to determine the absolute position of the epigenetic event given the underlying molecule sequence.[37-40]

Though there exist many approaches to nanochannel manipulation of DNA or chromatin, several in particular have been demonstrated to be applicable towards the direct detection and imaging of epigenetic marks on either DNA or chromatin. One such demonstrates the use of electrophoretic flow to drive

100kb-length DNA into 45nm straight channels. Using a series of micropillars in front of the channel entrance to disentangle the DNA, the molecules remain linearized due to the sub-persistence length diameter (50nm in DNA) of the nanochannels.[41] In this study, sequence specific nicking sites are used to label the DNA to which the optically measured length between labels can be correlated to a sequence-specific position due to enzyme recognition sites used. The genomic position of fluorescently tagged marks, in this case major histocompatability complexe (MHC) sites, can then be determined after piecing together fragments using overlapping regions as shown in ***figure 2.3*** below.[42]

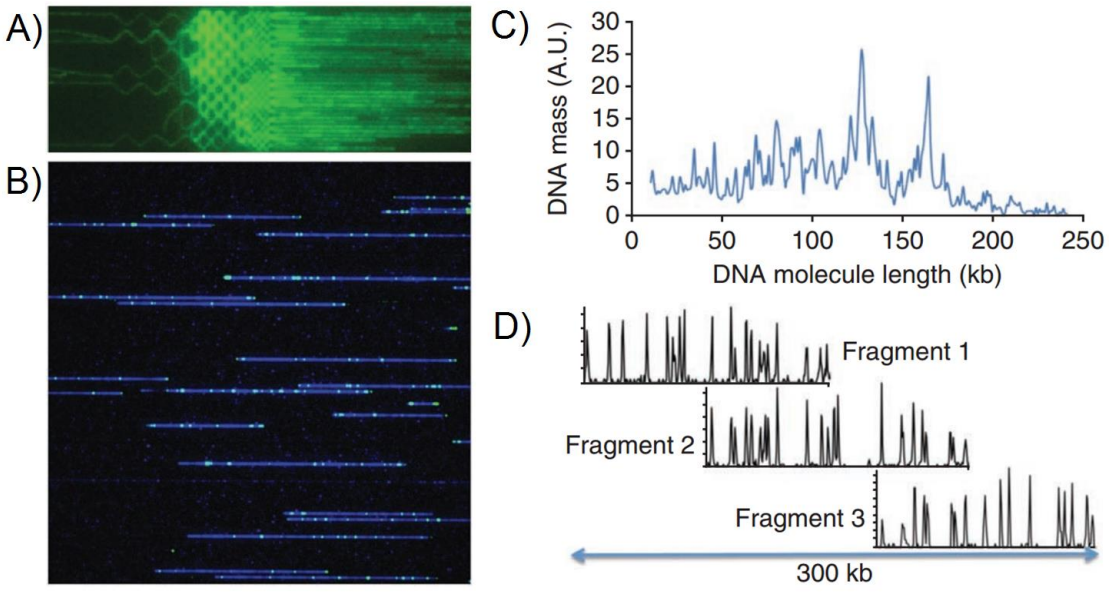

Figure 2.3 – Imaging and Mapping MHC sites on BACs in Nanofabricated Channels

(Modified from Lam E.T. et. al., 2012)

(A) DNA labeled with YOYO-1 intercalating dye demonstrates DNA's ability to enter nanochannels after micropillar facilitated disentanglement (Micropillars not highlighted). (B) Nick-labled MHC sites (green) on DNA backbone (blue) while the DNA is traversing the channel via electrophoretic flow. (C) DNA fragments are elongated to 85% of their theoretical length. Longest fragment size reaches up to 250kb. (D) MHC locations for each fragment can be mapped and stitched back together via overlapping regions.

While this is an elegant demonstration of high-resolution imaging based DNA mapping, the above nanochannels are far less compatible when applied to chromatin as protein-surface interactions can often cause clogging in electrophoretic flow based nanochannels. Furthermore, fixed nanochannel dimensions do not reflect the higher complexity of chromatin. For instance, chromatin flexibility and persistence lengths vary based on histone modifications and irregularities in nucleosome folding. An alternative approach to confine and elongate chromatin.[43] Demonstrated by Matsuoka T. and colleagues in 2012, ***figure 2.4***, capitalizes on the deformability PDMS to elongate chromatin by the action of constricting channel width.[44] By pulling axially on a 122nm diameter PDMS nanochannel, a combination of shear forces

and reduced entropic freedom was used to elongate λ-DNA to 97% of its contour length. This technique was then applied to diamidino phenylindole (DAPI) stained HeLa-GFP chromatin labeled with histone H4 acetylation (H4ac) antibody. The H4ac marks are then mapped to histone positions and the underlying DNA.

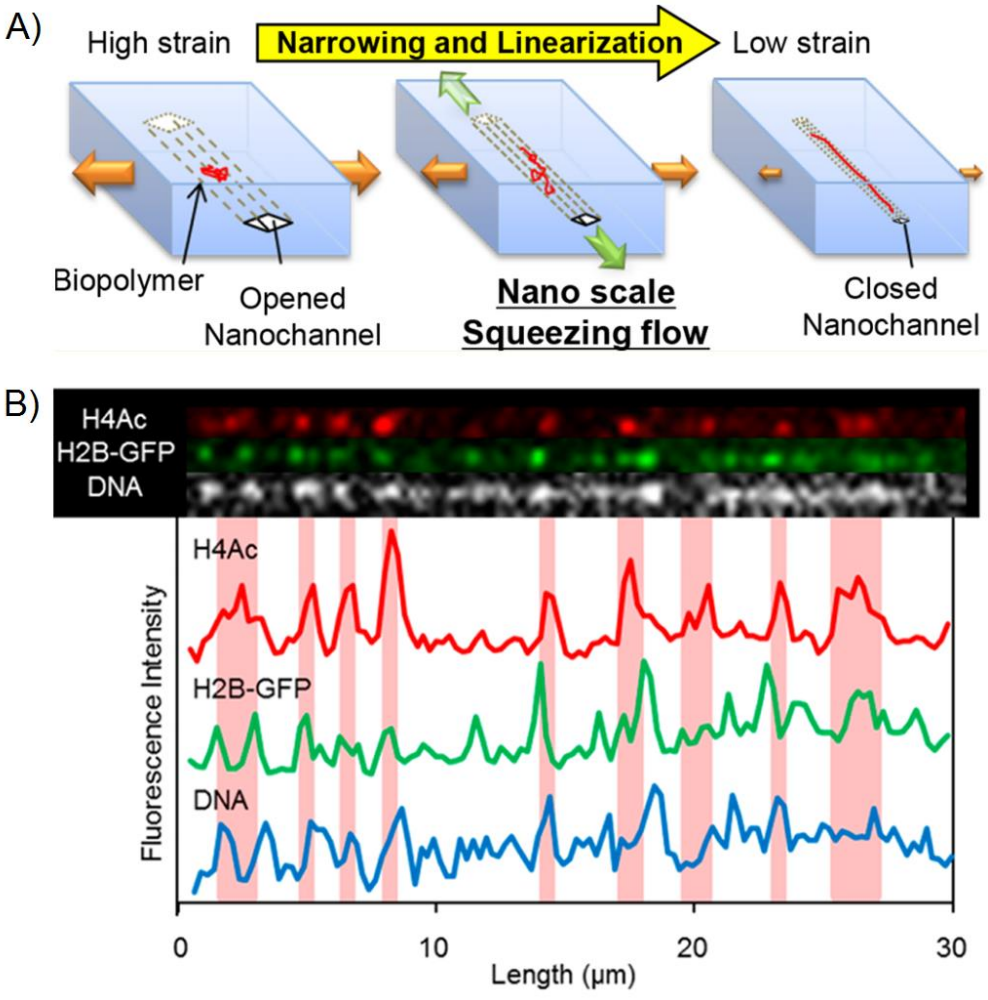

Figure 2.4 – PDMS Nanoconfinement Device for Chromatin Elongation and Epigenetic Mapping

(Modified from Matsuoka T. et. al., 2012)

(A) As the PDMS nanochannel is extended axially, the constricting channel width causes elongation of HeLa-GFP chromatin molecules within the channel. A proof of concept type experiment is shown in (B) where histone H4 acetylation labeled with Anti-H4 antibody and mapped to the position of green fluorescent protein (GFP) fluorescent nucleosomes and the underlying DNA stained by DAPI.

Authors in the above study observe that even when DNA is extended to over 80% of its contour length, less than 35% of the DNA will be fully extended. The majority of molecules contain a knot either in the form of a dumbbell at one of the distal end of the molecule or a kink proximal to the center. This suggests that reliable extension of chromatin may not be guaranteed. Hence, the advantages of high-throughput nanochannel techniques may be more suited for genetic or epigenetic analysis of bare DNA as opposed to chromatin.

2.3.4 Surface Immobilization Technologies

In contrast to nanochannel and confinement type technologies, surface immobilization technologies are not prone to non-specific binding induced channel clogging and therefore favorably suited to process and handle native chromatin. Techniques in surface immobilization of DNA date back to the mid-1990s with work from Schwartz D. C. and colleagues where the interface of a drying droplet was used to elongate λ-DNA on a silanized glass surface.[45-46] With restriction endonuclease BamHI, they created nick sites without perturbing the elongated orientation of the molecules in a proof-of-concept experiment towards using surface immobilization of as a tool for genetic mapping. Bensimon A. and colleagues showed a similar capacity to elongate DNA by slowly vertically lifting a silanized glass slide from a beaker containing DNA in solution.[47-49] This process, known as molecular combing, aligns molecules perpendicular to the edge of a receding liquid meniscus. As demonstrated by Labit H. et al in 2008, ***figure 2.5***, fluorescence in-situ hybridization (FISH) probes for specific sequences can be hybridized onto surface immobilized combed-DNA for genetic mapping. However, although this method has been applied in various instances of mapping DNA, it has yet to be applied to the mapping native euchromatin.

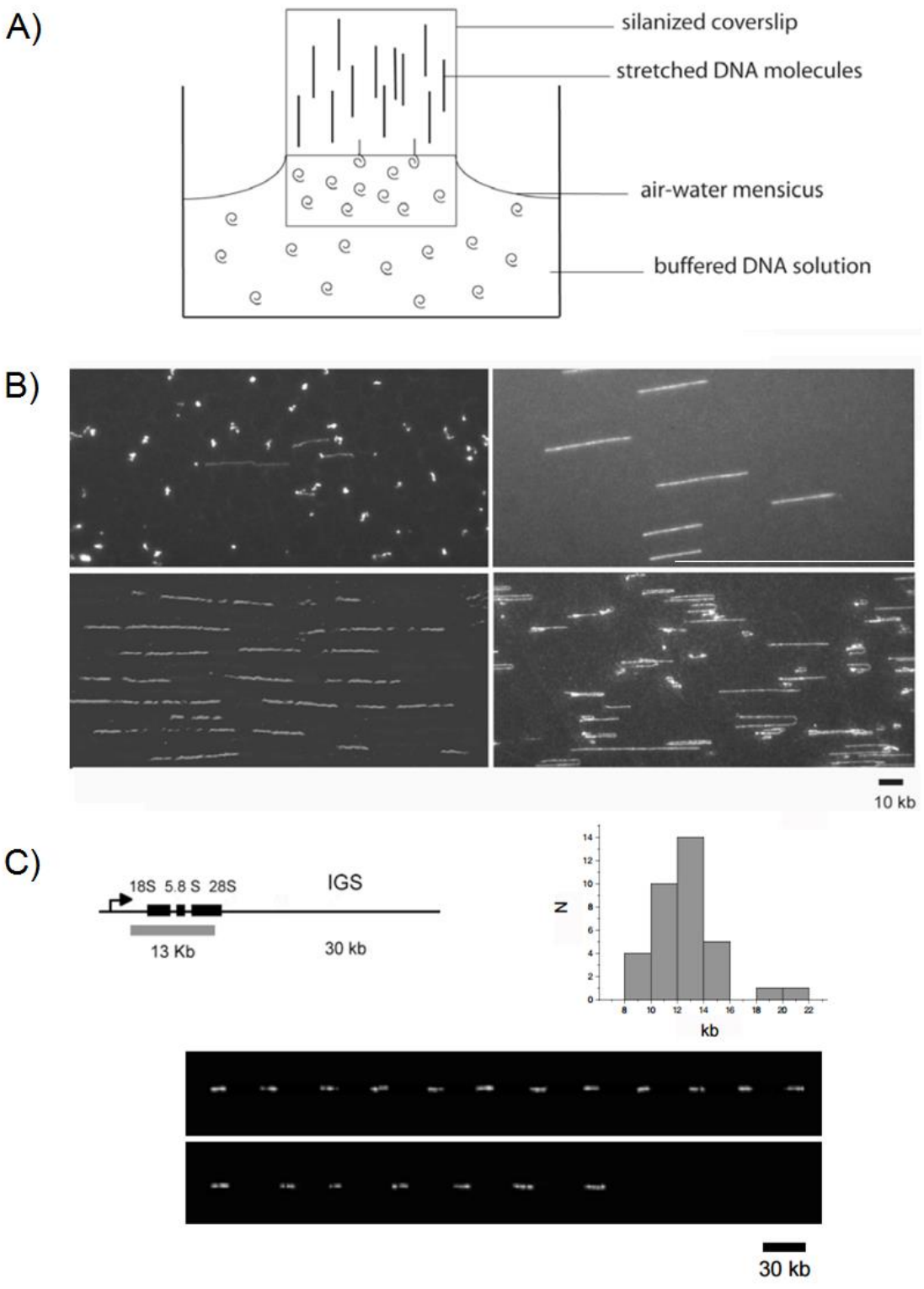
A)
silanized coverslip
stretched DNA molecules
air-water mensicus
buffered DNA solution
B)
10 kb
C)
18S 5.8 S 28S
IGS
13 Kb
30 kb
N
kb
30 kb

Figure 2.5 – FISH Labeling of λ-DNA Surface-Immobilized by Molecular Combing

(Modified from Labit H. et. al., 2008)

(A) DNA in solution is elongated at the liquid-to-air interface as a silanized glass coverslip is lifted from the solution at a constant velocity of 300 μm s^{-1}. (B) Micrographs of immobilized λ-DNA stained with YOYO-1 intercalating dye. The DNA is combed into coiled, kinked, looped, and fully elongated configurations depending on surface treatment. (C) Repeating unit labeled by rDNA FISH probes to demonstrate fluorescence mapping. The length of the probes measured in the micrographs corresponds to the known length of the probes.

Lastly, mapping has been demonstrated via non-immobilization based surface linearization. Using a lipid bilayer to generate a fluid surface, biotinylated DNA molecules can be end-tethered to select lipids. The resulting tethered molecule is free floating when placed solution but under hydrodynamic flow, the molecules can be pulled and elongated in the direction of flow. Coupling this principle with microfabricated barriers on the bilayer surface, molecules flowing downstream can be anchored along the barrier and linearized in a highly parallel fashion. This technique, known as DNA curtains, allows for a high throughput means of imaging DNA. Recently, the technique has been demonstrated in application with reconstituted chromatin to study nucleosome distribution as shown in ***figure 2.6***.[50]

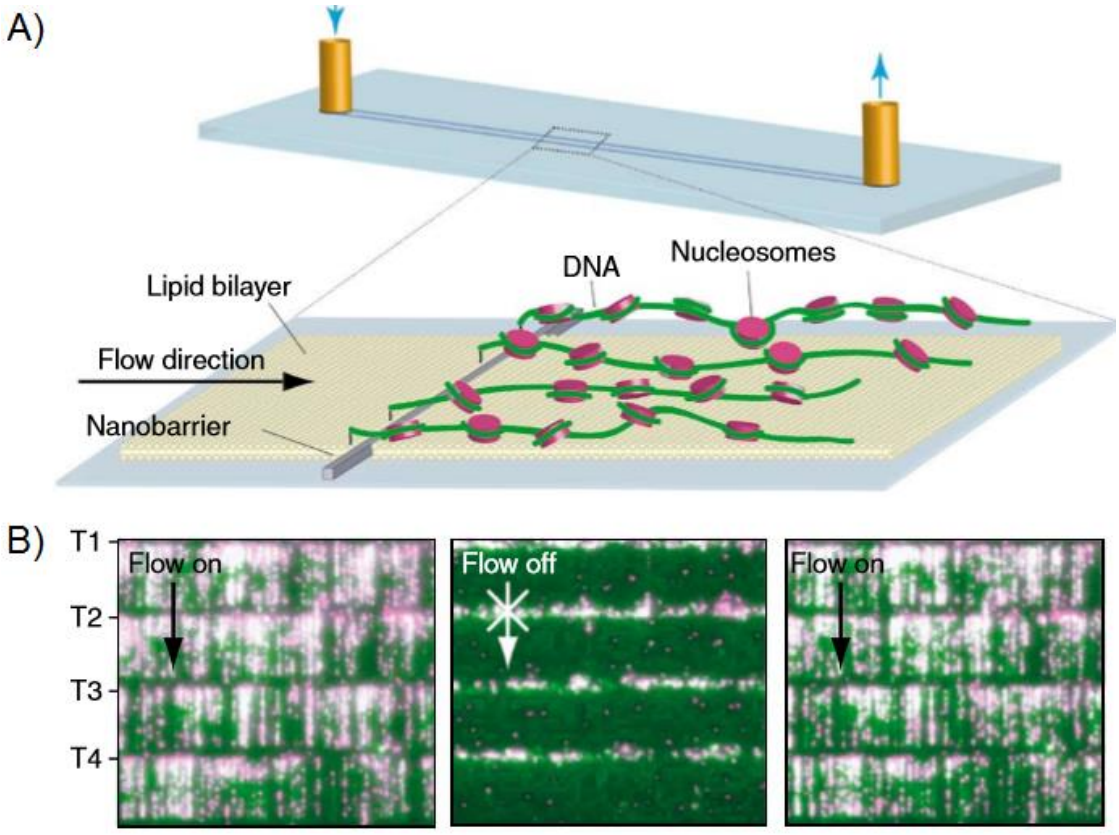

Figure 2.6 – Visualizing Nucleosomes on Reconstituted λ-DNA Using the Curtains Linearization Technique

(Adapted from Visnapuu M. L. et. al., 2009)

(A) Experimental overview of curtains linearization applied to chromatin. λ-DNA reconstituted chromatin is tethered to the lipid bilayer via a biotin-neutravidin interaction and hydrodynamic forces elongate the molecule in the direction of flow. (B)

Micrographs depicting chromatin molecules with quantum-dot labeled nucleosomes (pink) and YOYO-1 intercalating dye labeled DNA backbone (green).

Even though the authors above demonstrate the high degree of parallelization conferred by the curtains technique on reconstituted chromatin, there are inherent disadvantages to the technique when considering native chromatin. Akin to the obstacles faced by using tweezer based methods for chromatin elongation, techniques that require biotin conjugation are less suitable when the sequence of the DNA or chromatin molecule of interest is unknown.

2.4 CONCLUSION

Here, techniques used for elongation and linearization of DNA and chromatin as it is applied to genetic and epigenetic mapping have been outlined. Collectively, these techniques preliminarily demonstrate the potential of direct interrogation of the genome. Each category of technologies described confer a set of advantages over indirect analysis methods such as ChIP-seq or bisulfite sequencing, however, these are often niche improvements that come as tradeoffs to existing capabilities.

Tweezer based methods offer the highest degree of control over a molecule of interest, making it appropriate for studying sensitive mechanical and structural forces of DNA or chromatin, but suffers from extremely low throughput and inability to end-tether a nucleic acid molecule of unknown sequence. On the other hand, methods such as DNA curtains offer a much higher throughput, able to visualize thousands of molecules on a single device, yet still require a known sequence to end-tether the DNA or chromatin strand. Nanochannel, nanopores, and nanoconfinement can load a DNA or chromatin molecule of any sequence, but each suffers from its own shortcomings. Nanochannels and nanopores are prone to clogging resulting from non-specific protein interactions whereas fully extending molecules via nanoconfinement is a Poisson dependent process.

Furthermore, most techniques have only been demonstrated with DNA rather than chromatin. Of the techniques that have demonstrated their application with chromatin, fewer still have done so with native chromatin as opposed to reconstituted chromatin. Current challenges in working with native chromatin are numerous, including the extraction of long native chromatin strands and high throughput sequence independent elongation. In the chapter following, we expand upon molecular combing as a basis for elongation of native chromatin and imaging epigenetic marks via fluorescence labeling.

2.5 REFERENCES

1. "The Human Genome Project Completion" Genome.gov. National Human Genome Research Institute. 30 October 2010. Accessed 01 February 2017
2. T. J. Richmond, C. A. Davey, Nature, 2003, 423, 145-150
3. M. Vignali, J. L. Workman, Nature Structural Biology, 1998, 5, 1025-1028
4. W. Reik., Nature, 2007, 447, 425-432
5. J. C. Eissenberg, S. CR. Elgin, Chromosomes and Expression Mechanisms, 2014, eLS, 1-9
6. S. I. S. Grewal, S. Jia, Nature, 2007, 8, 35-46
7. O. L. Miller Jr., B. R. Beatty, Science, 1969, 164(3882), 955-957
8. A. Portela, M. Esteller, Nature, 2010, 28, 1057–1068
9. H. Y. Zoghbi, A. L. Beaudet, Cold Spring Harbor Persp. in Bio., 2016, 1-29
10. A. P. Feinberg, Nature, 2007, 447, 433-440
11. F. Antequera, Cellular and Molecular Life Sciences, 2003, 60, 1647-1658
12. M. Esteller, Nature, 2002, 21, 5427-5440

13. M. G. Zulueta, C. M. Bender, A. S. Yang, T. D. Nguyen, R. W. Beart, J. M. V. Tornout, P. A. Jones, Cancer Research, 1995, 55, 4531-4535
14. J. C. Lin, S. Jeong, G. Liang, D. Takai, M. Fatemi, Y. C. Tsai, G. Egger, E. N. Gal-Yan, P. A. Jones, Cell, 2014, 12, 432-444
15. F. Magdinier, L. M. Billard, G. Wittmann, L. Frappart, M. Benchaib, G. M. Lenoir, J. F. Guerin, R. Dante, FASEB Journal, 2000, 1585-1594
16. C. A. Aguilar, H. G. Craighead, Nature Nanotechnology, 2013, 8, 709–718
17. L. P. O'Neill, M. D. VerMilyea, B. M. Turner, Nature Genetics, 2006, 38, 835 - 841
18. M. C. Miller, G. V. Doyle, L. W. M. M. Terstappen, J. of Oncology, 2010, 617421, 1-8
19. M. L. Sakin, Y. Ebenstein, Current Opinion in Biotech., 2013, 24, 690-698
20. S. B. Smith, C. Bustamante, Science, 1992, 258(5085), 1122-1126
21. S. B. Smith, Y. Cui, C. Bustamante, Science, 1996, 271(5250), 795-799
22. U. Bockelmann, P. Thomen, B. E. Roulet, V. Viasnoff, F. Heslot, Biophysics Journal, 2002, 82, 1537-1553
23. J. J. Hayes, J. C. Hansen, PNAS, 2002, 99(4), 1752-1754
24. B. Fierz, ACS Chem. Bio., 2016, 11, 609-620
25. H. Oana, K. Nishikawa, H. Matsuhara, A. Yamamoto, T. G. Yamamoto, T. Haraguchi, Y. Hiraoka, M. Washizu, Lab Chip, 2014, 696-704

26. Yanagi, T. Ishida, K. Fujisaki, K. I. Takeda, Scientific Reports, 2015, 5, 14656, 1-13
27. E. Paulechka, T. A. Wassenaar, K. Kroenlein, A. Kazakov, A. Smolyanitsky, Nanoscale, 2016, 8, 1861-1867
28. S. J. Heerema, C. Dekker, Nature Nanotechnology, 2016, 11, 127-136
29. M. Akeson, D. Branton, J. J. Kasianowicz, E. Brandin, D. W. Deamer, Biophysics Journal, 1999, 77, 3227-3233
30. L. Mereuta, M. Roy, A. Asandei, J. K. Lee, Y. Park, I. Andricioaei, T. Luchian, Scientific Reports, 2014, 4, 3885, 1-11
31. Shim, G. I. Humphreys, B. M. Venkatesan, J. M. Munz, X. Zou, C. Sathe, K. Schulten, F. Kosari, A. M. Nardulli, G. Vasmatzis, R. Bashir, Scientific Reports, 2013, 3, 1389, 1-8
32. F. Haque, J. Li, H. C. Wu, X. J. Liang, P. Guo, Nano Today, 2013, 8(1), 56-74
33. R. E. Gyurcsanyi, Trends in Anal Chem, 2008, 27(7), 627-639
34. G. V. Soni, C. Dekker, Nano Letters, 2012, 12(6), 3180-3186
35. S. Agah, M. Zheng, M. Pasquali, A. B. Kolomeisky, J. of Physics D: Applied Physics, 2016, 46, 413001 (14pp), 1-14
36. G. F. Schneider, Q. Xu, S. Hage, S. Luik, J. N. H. Spoor, S. Malladi, H. Zandbergen, C. Dekker, Nature Communications, 2013, 4, 2619, 1-7

37. S. K. Das, M. D. Austin, M. C. Akana, P. Deshpande, H. Cao, M. Xiao, Nucleic Acids Research, 2010, 1-8

38. M. Xiao, A. Phong, C. Ha, T. F. Chan, D. Cai, L. Leung, E. Wan, A. L. Kistler, J. L. DeRisi, P. R. Selvin, P. Y. Kwok, Nucleic Acids Research, 2007, 35(3), e16, 1-12

39. M. Baday, A. Cravens. A. Hastie, H. Kim, D. E. Kudeki, P. Y. Kwok, M. Xiao, P. R. Selvin, Nano Letters, 2012, 12, 3861-3866

40. A. R. Hastie, L. Dong, A. Smith, J. Finklestein, E. T. Lam, N. Huo, H. Cao, P. Y. Kwok, K. R. Deal, J. Dvorak, M. C. Luo, Y. Gu, M. Xiao, PLOS One, 2013, 8(2), e55864, 110

41. E. Y. Chan, N. M. Goncalves, R. A. Haeusler, A. J. Hatch, J. W. Larson, A. M. Maletta, G. R. Yantz, E. D. Carstea, M. Fuchs, G. G. Wong, S. R. Gullans, R. Gilmanshin, Genome Research, 2004, 14, 1137-1146

42. E. T. Lam, A. Hastie, C. Lin, D. Ehrlich, S. K. Das, M. D. Austin, P. Deshpande, H. Cao, N. Nagarajan, M. Xiao, P. Y. Kwok, Nature Biotech, 2012, 30(8), 771-777

43. S. F. Lim, A. Karpusenko, J. J. Sakon, J. A. Hook, T. A. Lamar, R. Reihn, Biomicrofluidics, 2011, 5, 034106, 1-8

44. T. Matsuoka, B. C. Kim, J. Huang, N. J. Douville, M. D. Thouless, S. Takayama, Nano Letters, 2012, 12, 6480-6484

45. J. Jing, Z. Lai, C. Aston, J. Lin, D. J. Carucci, M. J. Gardner, B. Mishra, T. S. Anantharaman, H. Tettelin, L. M. Cummings, S. L. Hoffman, J. C. Venter, D. C. Schwartz, Genome Research, 1999, 9, 175-181
46. Z. Lai, J. Jing, C. Aston, V. Clarke, J. Apodaca, E. T. Dimalanta, D. J. Carucci, M. J. Gardner, B. Mishra, T. S. Anantharaman, S. Paxia, S. L. Hoffman, J. C. Venter, E. D. Huff, D. C. Schwartz, Nature Genetics, 1999, 23, 309-313
47. A. Bensimon, A. Smion, A. Chiffaudel, V. Croquette, F. Heslot, D. Bensimon, Science, 1994, 265(30), 2096-2098
48. J. C. Allemand, D. Bensimon, L. Jullien, A. Bensimon, V. Croquette, Biophysics Journal, 1997, 73, 2064-2070
49. J. Herrick, A. Bensimon, Biochimie, 1999, 81, 859-871
50. M. L. Visnapuu, E. C. Greene, Nature Struc. Mol. Biol., 2009, 16(10), 1056-1062

CHAPTER 3

IMAGING EPIGENETIC MARKS ON DNA AND CHROMATIN ARRAYS

3.1 INTRODUCTION

In the previous chapter, we reviewed current standard techniques in epigenetic mapping as well as emerging techniques based on high resolution imaging of fluorescently labeled marks to directly interrogate the epigenetic landscape of a DNA or chromatin molecule. Of the techniques reviewed, few were suitable for native chromatin due to the complexity of extracting long, intact mammalian chromatin molecules of unknown sequence and then elongating it with high throughput methods. Here, we discuss the process of ultra-long native chromatin extraction from mammalian cells and subsequent elongation, immobilization, and finally, epigenetic labeling and imaging of the native chromatin.

Techniques such as fluorescence in-situ hybridization (FISH) has long been used to fluorescently label target sequences on chromosome molecules.[1-3] The basic working principle of FISH is to use a target specific DNA probe that has been fluorescently labeled through processes such as nick translation,

where DNA polymerases are used to shift a DNase induced nick to allow for the insertion of fluorescent tags. This fluorescent DNA probe is then hybridized to a target sequence within a chromosome or segment of DNA and marks it with the specified fluorescence signature. Numerous applications have arisen out of FISH such as diagnosing chromosomal aberrations, interphase characterization for cytogenetics, monitoring disease progression, identification of translocation break points, gene mapping, and even painting whole chromosomes.[4-16] With the introduction of molecular combing by Bensimon et. al. in 1990s, molecular combing has improved the resolution of FISH based mapping down to 1 kb.[17-18]

More recently, the process of molecular combing has been further optimized with the advent of microfabrication. One such improvement involves the use of photolithography to generate a pattern from which an array of microwells can be mold-casted. This microwell array can then serve in controlling the positioning of molecules being combed across the surface to achieve an ordered array. This ordered array patterning lends to improved image analysis and potential automation with regular spacing between molecules. First showcased with naked DNA in 2005 by Guan J. et. al., DNA can not only be combed into aligned and ordered arrays, but they can be transferred onto a separate substrate via in a technique known as transfer printing.[19] In transfer printing,

DNA is combed on a polymer stamp that contains the patterned microwells and before the stamp face is contacted to a silanized surface. The method is simple to perform requiring very little equipment set up and robust, serving as an effective means of elongating DNA.[20-21] For these reason, our group has previously used this method to demonstrated epigenetic imaging of methylation sites on bacteria lambda phage DNA as shown in ***figure 3.1*** below.[22] Here, methyl-CpG-binding domain (MBD) peptide that has been fluorescently labeled with Alexa Fluor 488 dye is used to tag CpG dinucleotides for identification upon fluorescence imaging.

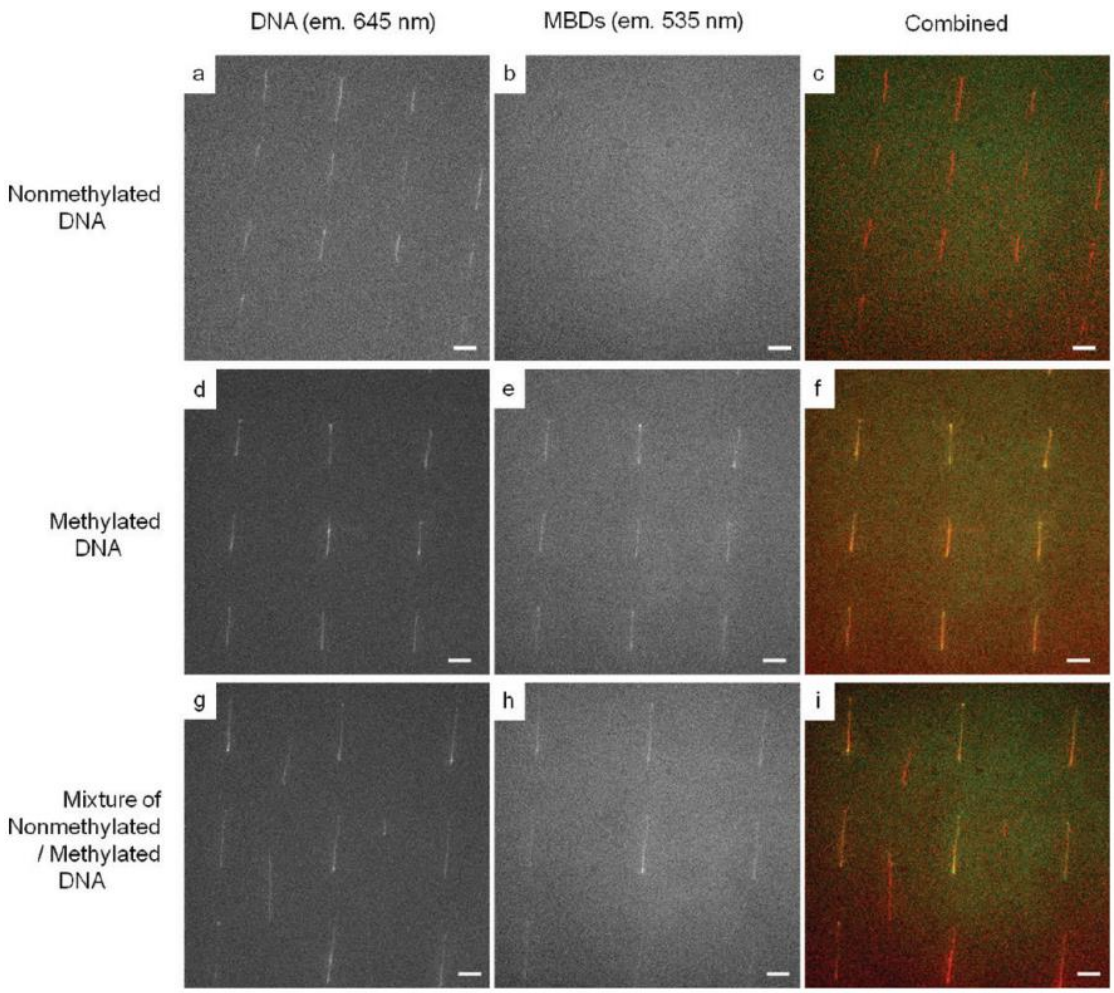

Figure 3.1 – Imaging DNA Methylation with MBD Peptide on DNA Arrays

(Adapted from Cerf A. et. al., 2011)

Molecular combing in junction with contact transfer printing is used to elongate λ-DNA molecules and immobilize them on silanized glass cover slip. (Row 1) Unmethylated DNA is used as a control and shows no non-specific MBD peptide binding. (Row 2) Methylated DNA is combed and visible under the wavelengths exciting the MBD labels. The combined image shows methylation sites (red) in addition to the naked DNA (green). (Row 3) Shows the combed molecules from a 1:1 mixture of

methylated and unmethylated DNA. Expectedly, some molecules are not present when exciting under 475nm (emission 535nm) due to the lack of methylation sites on those molecules.

Coupled with high-resolution imaging such as total internal reflect (TIRF) microscopy or localization microscopy, molecular combed DNA labeled with fluorescent tags may reach sufficient spatial resolution for direct imaging-based DNA methylation mapping.[23-25] To then demonstrate the potential for this technique to serve as a high-throughput sequence independent platform for epigenetic mapping of chromatin, we transposed these methods to native human chromatin derived from cancer cells and then labelled for various histone modifications.

3.2 MATERIALS AND METHODS

3.2.1 Polymer Microwell Stamp Fabrication

Microwell patterns were etched in standard 4-inch silicon wafers with a mask containing 9 identical stamps arranged in a 3 x 3 grid aligned to the center of the mask. Each stamp contains arrays of 5μm and 8μm diameter circles spaced

in with varying distances between wells ranging 15μm or 20μm apart depending on the region of the stamp. Since combing selects for molecules within certain size ranges, these different patterns maximize the likelihood of combing success for a wider range of molecule sizes. Using the ABM Near-UV Vacuum Contact Aligner, the pattern from the mask was transferred to the resist spun onto the silicon wafer. The resist was then stripped and the wafer was etched in a Unaxis SLR 770 Deep Si Etcher for 15 cycles. Height of the resulting pillars on the silicon master was measured using a P-10 profilometer. These posts serve as a negative mold for polymer mold casting the microwell array polymer stamps. Polydimethylsiloxane (PDMS) was mixed at a ratio of 10:1 and poured onto the silicon master until a 0.5cm thick layer of PDMS was achieved. The PDMS was then baked in a closed oven at 80C for 3 hours. Once fully cured, the PDMS slab was peeled off from the wafer and the combing stamps were excised from the slab with an X-Acto knife as rectangular stamps measuring approximately 1.4cm x 1.8cm.

3.2.2 Cell Culture

M0-91 acute myelogenous leukemia cell line cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal bovine serum

(FBS), 1% (vl) Pen-Strep, 1% (vl) Non-essential Amino Acids (NEAA), 1% (vl) Sodium Pyruvate, 1% (vl) L-Glutamine 200mM, 2.5% (vl) 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), and 0.1% (vl) of 1% 2-mercaptoethanol (β-ME). Media was filtered with a 20μm filter prior to use. Cells were cultured in a Sheldon 2350-T incubator set at 37C and 5% CO_2 until 70%-80% confluency. Cells were split taking aliquots of the M0-91 cells grown in culture and adding them directly to fresh media at a 1:50 dilution.

HeLa wild type (HeLa-WT) cervical cancer cell lines cells were cultured in identical media as the M0-91 cells for all chromatin combing experiments and also allowed to propagate to 70-80% confluency. To passage the cells, cell media was aspirated and the cells were washed with 1xPBS that has been set to 37C in a water bath. The PBS was aspirated out and 6ml of 0.25% Trypsin was added to the T75 culture flask and incubated with the cells at room temperature for 5 minutes. Cells were gently tapped to release them from surface adhesion and trypsin was neutralized with equal volumes cell media. The cell suspension was then transferred to a 15ml centrifuge tube and spun at 1000g for 5 minutes. The supernatant was discarded and the remaining cell pellet was resuspended in 5ml of fresh media. The resuspended cells were then aliquoted into new T75 flasks at a dilution of 1:50.

3.2.3 Ultra-Long Native Chromatin Extraction

Prior to extraction, 30ml of chromatin extraction buffer was prepared on ice containing 10.2ml 1M sucrose, 0.45ml 1M Tris (base), 0.45ml 1M NaCl, 1.8ml 1M KCl, 0.3ml 1M $CaCl_2$, 37.5ul 200mM phenylmethylsulfonyl fluoride (PMSF), 30ul 100% β-ME, 15ul spermidine, and 4.5ul spermine. Each buffer must be pH adjusted to pH 8 prior to mixing into chromatin extraction buffer. After mixing all components together. Separately, lysis buffer containing 1% Triton-X in 1x phosphate buffered saline (PBS) is supplemented with 1 tablet of c0mplete Ultra Mini EDTA-free protease inhibitor (Roche) for every 50ml of lysis buffer.

The first step of the extraction procedure is to remove the cells from the T75 flasks in which they are grown. Since M0-91 cells are non-adherent cells, no trypsin is necessary. Cells are transferred into 15ml BD-plastic capped test tubes and spun at 1000g for 10 minutes in an Allegra 21R centrifuge. The supernatant from the centrifuged test tubes are aspirated and discarded. The remaining cell pellet is resuspended in 5ml of lysis buffer and physically shaken to break up the cell before placing the eppendorf on ice for 5 minutes. The cell suspension was then transferred to a 7ml glass dounce homogenizer and homogenized for 12 strokes, with an up and down motion of the dounce hammer counting as 2

strokes. After homogenization, the lysis buffer and lysed cell suspension within the dounce is transferred aliquoted into 1.5ml eppendorf tubes with a maximum volume of 1ml per tube. The eppendorf tubes containing the M0-91 chromatin, cellular debris, and lysis buffer was then placed in an IEC Micromax centrifuge and spun at 10,000g for 10minutes. The supernatant from each eppendorf tube was carefully aspirated out with a P-200 micropipette and the remaining pellet, which is the M0-91 chromatin, was resuspended in chromatin extraction buffer prepared above.

The chromatin extraction process relies on a balance between chemical and mechanical lysis to extract chromatin while maintaining nucleosome integrity. 1% Triton-X serves to lyse the cellular membrane while keeping the nuclear envelope intact. Chemicals able to dissolve the nuclear envelope are far too harsh for maintaining the nucleosomes, and thus the mechanical shear forces generated by the dounce homogenizer is meant to create tears in the nuclear envelope. Thus, the chromatin must be given time to diffuse out of the ripped nuclear envelopes. This process can take up to two weeks for ultra-long chromatin. Alternatively, if shorter chromatin fragments are desired, the diffusion time can be shortened by digesting the chromatin contained within the torn nuclear envelopes with micrococcal nuclease for until a desired size range is yielded and then heat inactivating the enzyme.

3.2.4 Histone Antibody Labeling

1:200 dilution of monoclonal mouse histone H3 primary antibody MABI 0301 (Active Motif) was incubated with 2ug of native chromatin for 1 hour at room temperature. AlexaFluor647-tagged goat anti-mouse IgG secondary antibody (Invitrogen) was then incubated with the sample at 1:200 dilution for an additional 30 minutes in the absence of light. This process was repeated for histone H3K9me2 labeling using H3K9me2 rabbit antibody (Active Motif) and AlexaFluor488-tagged anti-rabbit IgG secondary antibody (Invitrogen). The histone labeled chromatin samples were then incubated with 1:10,000 YOYO-1 intercalating dye (Invitrogen) or 1:200 DAPI (4',6-Diamidino-2-Phenylindole, Dihydrochloride) (ThermoFisher Scientific) for 1 hour at room temperature also in the absence of light.

3.2.5 Chromatin Arrays Molecular Combing

PDMS stamps containing the microwell array patterns were fabricated as described in previous sections of the methods and used within minutes of

peeling the stamp off of the master as to prevent dust adsorbance onto the surface. All combing was performed at room temperature under the absence of light to prevent potential photobleaching of any fluorescent labels on the chromatin. The setup for the combing process involves a mechanically translating stage that is controlled by a Newport 861 motion controller and an adjustable lever arm attached to a glass slide.

To begin, a PDMS stamp is laid flat on the surface of the translating stage while under no motion. A 15ul droplet containing chromatin at 200ng ul^{-1} is loaded onto the surface of the stamp towards one of the axial ends of the stamp via a P-200 micropipette. The stamp orientation is determined by the 1.4 x 1.8 dimension with the longer axis aligned to the direction of combing. One half of a 180µm thick glass cover slip with the dimensions of a standard microscope slide is then adhered onto an adjustable lever arm. This lever arm is placed along the opposing end of the direction of combing and the glass cover slip is carefully manually lowered onto the PDMS surface until the chromatin droplet has made contact with the glass. The cover slip is then further lowered until there is a sub-1mm gap remaining between the stamp and the cover slip, allowing the droplet containing the chromatin sample to spread out and span the entire width of the stamp. When these processes have been completed, the sample is considered to be loaded.

An illustration of the operational procedures is shown in ***figure 3.2*** below. Briefly, once the sample of native chromatin has been loaded, the stage upon which the PDMS stamp sits is mechanically translated away from the position of the lever arm at a constant velocity of 20μm s^{-1} until the cover slip, that once overlapped the stamp, is no longer overlapping the stamp. The surface charge of the glass coverslip allows the sample to more favorably reside on the coverslip surface rather than the hydrophobic PDMS surface throughout the entire combing process. As the stamp is translated away from the cover slip that is held by the lever arm, the liquid contact line is swept over the entire PDMS stamp face. As the liquid-air interface passes over the patterned microwells, a combination of its velocity, the liquid contact angle, and local evaporation causes chromatin molecules within the solution to become pinned within the microwell. The pinned molecule is then elongated as the interface is pulled across the microwell and remaining is the chromatin molecule in a linearized configuration on the stamp surface. Once combing was completed, the chromatin arrays arrested on the surface of the PDMS stamp was then transferred onto a Fisherbrand 22x22-1 glass cover slip (Fisher Scientific) following the Microcontact transfer printing protocols outlined in the following section.

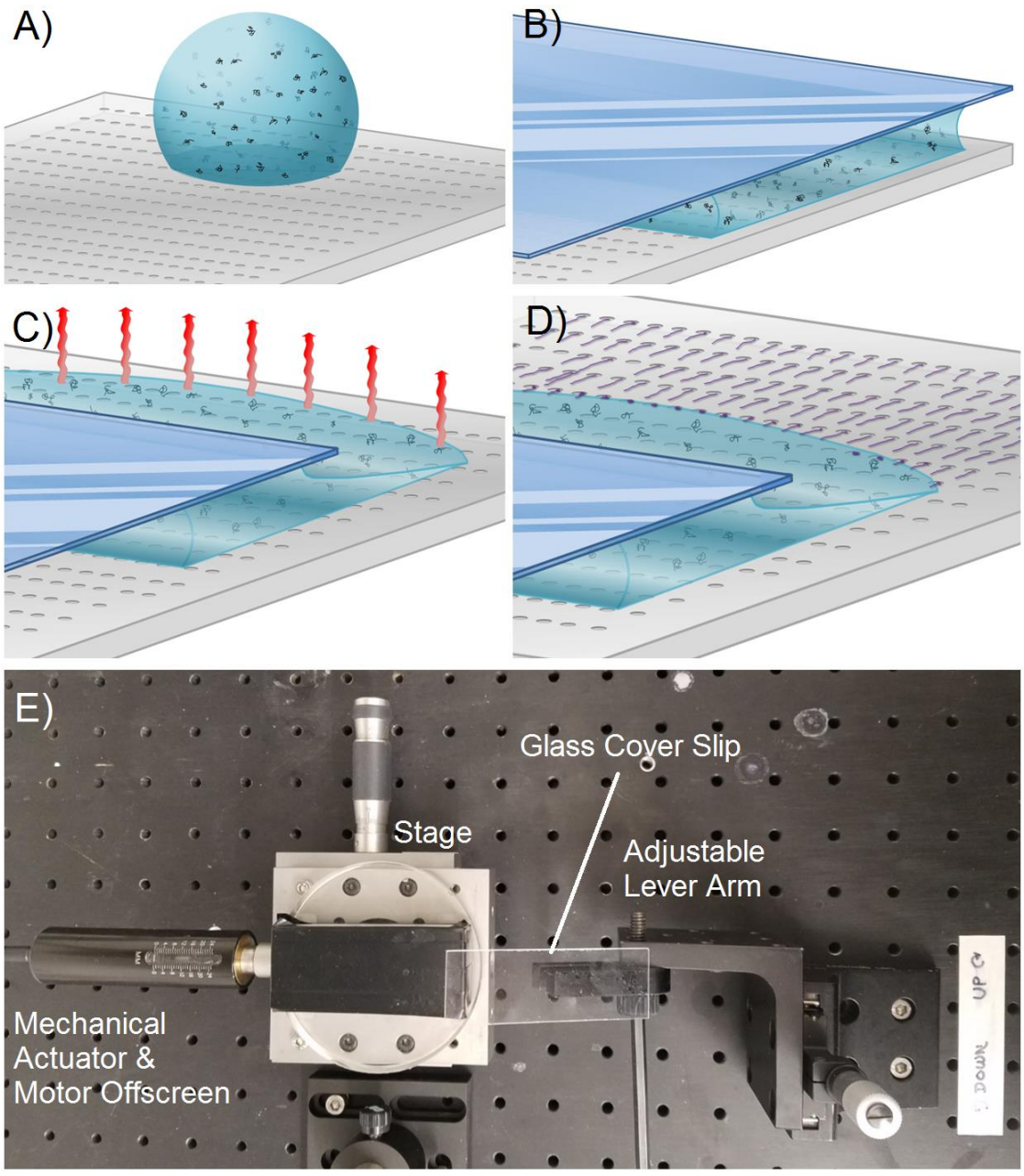

Figure 3.2 – Chromatin Molecular Combing Operational Procedure and Experimental Setup

(A) A 15ul droplet of sample containing chromatin (black lines floating in the liquid) is loaded onto a PDMS 0.5cm thick PDMS stamp via micropipette. (B) A glass slide is then lowered onto the droplet, sandwiching the droplet and spreading the liquid to span a wider area of coverage. (C) As the stage supporting the polymer stamp is translated away from the position of the fixed glass cover slip, the sample travels with the hydrophilic glass surface instead of the hydrophobic stamp surface. This causes the droplet to be dragged over the patterned microarrays on the stamp and during this process, a combination of forces including edge evaporation (red arrows) causes the chromatin molecules to be pinned within microwells. (D) As the fluid front passes over the microwell array, pinned chromatin molecules are pulled and elongated onto the stamp surface. (E) Top-down view of the experimental setup on an optical table. Here, the adjustable lever arm can be rotated or locked with a hex key and the PDMS stamp would be loaded onto the black colored slab on the stage. Z-position of both the stage and lever arm can be adjusted.

3.2.6 Chromatin Microcontact Transfer Printing

Platinum Line 24mm x 60mm cover glass was sonicated for 4 cycles alternating between immersed in 100% ethanol and 1M KOH beginning with KOH and ending with ethanol cycle. The cover glass was kept in ethanol until use, and washed again with 100% ethanol and dried with bone-dry grade nitrogen gas.

Cleaned cover slips were treated for 5 minutes with oxygen plasma using a Harrick Plasma Cleaner. After oxygen plasma treatment, cover slips were immediately submerged in a beaker containing 1% APTES, 5% water, and 94% ethanol. After one hour, cover slips were removed from the beaker, gently rinsed with 100% ethanol first and then 100% water before being dried with nitrogen gas. Finally, the now silanized glass cover slips were placed on a hot plate set to 150C for 10 minutes. To microcontact transfer print the chromosome arrays, one edge of the silanized surface of the glass cover slip was contacted with the PDMS stamp containing the post-combed chromatin arrays. The cover slip was then carefully laid onto the cover slip by motion of hinged door as to prevent air bubble pockets. Full contact between the stamp and cover glass was visually confirmed and gentle pressure was applied to the marriage for 60 seconds before subsequently peeling the stamp off of the cover glass.

3.2.7 Chromatin Fragment Size Characterization

Chromatin extractions were characterized by gel electrophoresis after micrococcal nuclease digestion and removing histones via phenol chloroform extraction and ethanol precipitation. This was done to validate nucleosome integrity post-extraction. 300ul aliquots of chromatin obtained from the

chromatin extraction protocol outlined in above sections were digested with 1ul of 1:10 diluted micrococcal nuclease (2000 gel units) for times ranging from 5 minutes to 30 minutes. Separate tubes were digested for different times spanning 5 minute intervals per tube and enzymatic reactions were then neutralized with 35ul of 0.5M EDTA. Digested chromatin samples were then supplemented with 100ul 1%SDS, 100% neutral phenol, and 100% chloroform. The mixture was vortexed vigorously for 30 seconds and then spun at 10,000g for 5 minutes. The upper phase was carefully transferred into fresh 1.5ml eppendorf tubes. Equal volumes of 100% ethanol and 3.5ul of 3M NaOAc was added to the tube. The tube was centrifuged at 14,000 rpm in an Eppendorf 5424 centrifuge for 30 minutes while set at 4C. The supernatant was then removed and the pellet was gently washed with 70% ethanol in water. Finally, the washing solution was also aspirated out and the pellet was allowed to air dry for 1 hour before being resuspended in 300ul of 1x Tris-EDTA (TE) buffer. 500ng of the DNA was then ran on a 1.8% agarose gel at a constant voltage of 80V for 40 minutes. The gel was stained with 1:10,000 SYBR-Gold dye (ThermoFisher Scientific) on a rocking platform for 1 hour and imaged on a Kodak EDAS-290 UV lamp. Atomic force microscopy (AFM) characterization was performed with a Dimension 3000-1 AFM (Digital Instruments) using tapping mode at 80Hz and 512-line resolution.

3.3 RESULTS

3.3.1 Combing Characterization with Lambda Phage DNA

The protocol for chromatin molecular combing and microcontact transfer printing was developed and optimized on λ-DNA due to it being robust and easy to purchase. Lambda DNA contains 48,502 base pairs (bp) and with the theoretical length of each bp being 0.3nm, lambda should extend 14.55μm in length when fully elongated.[26] Lambda DNA stained with YOYO-1 intercalating dye was combed using the microwell pattern developed. Combing λ-DNA yielded a few interesting results. Firstly, an ideal concentration of DNA in solution prior to combing was determined to be 10ng to 30ng ul^{-1}. Much higher concentrations result in multiple fibers of DNA being bundled together and an inability to resolve single molecules. This decreases the accuracy of any potential genomic or epigenomic mapping applications as it becomes more difficult to determine which strand an observed mark of interest is originating from. Next, out of the various well sizes on a single stamp patterned area, multiple well sizes were observed to be able to trap and elongate λ-DNA during the combing process, however microwells spaced less 5μm apart and

microwells under 10µm in diameter were observed to exhibit poor trapping efficiencies. Examples of these phenomena are shown in ***figure 3.3*** below.

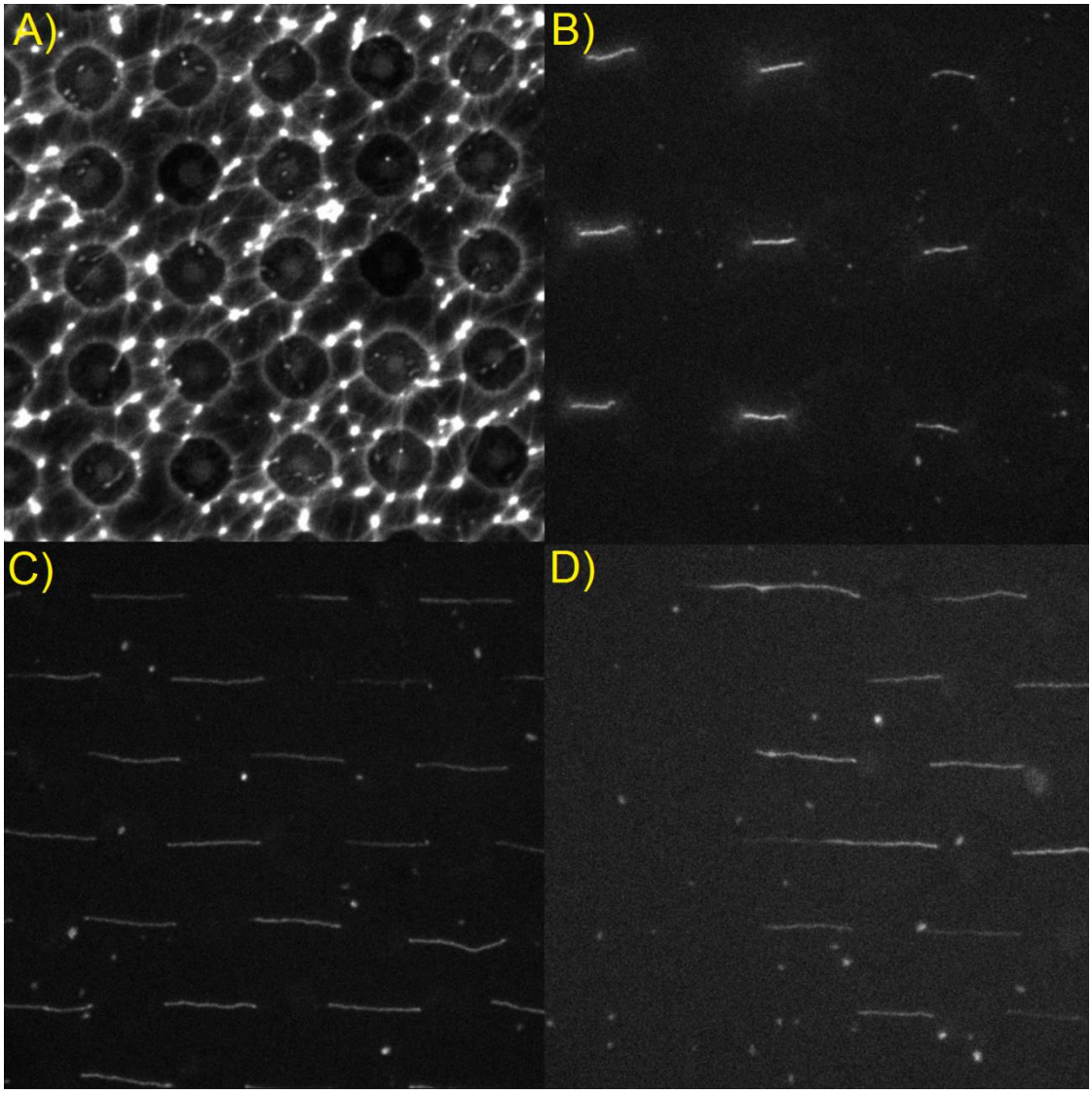

Figure 3.3 – Arrays of Lambda Phage DNA Molecular Combing Micrographs

(A) 2,000ng ul^{-1} λ-DNA combed on a PDMS stamp yields significant bundling and intertwining DNA molecules. (B) 25μm well diameter sizes with 10μm separation distance successfully combs and elongates λ-DNA as well as (C) 10μm well diameter sizes with a 15μm spacing. (D) A few molecules at the end of an array are elongated past the contour length of 14.55μm extending up to maximum observed length of 31.36μm. (Combing direction from right to left.)

Another observation of interest is that molecules at the end of the array can extend to longer lengths than 14.55μm. This suggests either stretching beyond the contour length or that there are multiple molecules intertwined and the observed strand is not a single molecule.[26] The former is entirely a possibility as due to the motion of the liquid front passing over the combing wells. In brief, the liquid front contours around a microwell until the meniscus force exceeds the pinning force and the liquid front snaps for a short distance and is caught on the next microwell as it repeats the contouring process. At the end of the array, the final snapping of the liquid front is not caught by another well and thus the snapping speed may be faster than throughout the rest of the array thereby stretching the λ-DNA molecule beyond its 14.5μm contour length. The alternative possibility of multiple molecules being combed in each well is not mutually exclusive with regards to the first scenario. To verify that chromatin molecules combed were single molecules rather than ropes, AFM was used to

characterize nucleosome and DNA backbone height. Using Image-J, the average measured λ-DNA strand length was 12.95μm (n=25) corresponding to an 89% extension, which is comparable to molecular combing directly on silanized glass.

3.3.2 Native Chromatin Extraction Characterization

Nucleosome integrity of native chromatin extractions were first verified with micrococcal nuclease digestion followed by gel electrophoresis to check for the presence of banding patterns. The micrococcal enzyme is a non-site specific endonuclease that cleaves DNA at random points. With chromatin, the enzyme will cleave around nucleosomes but are physically unable to access the DNA bound wrapped around a nucleosome whereas with bare DNA, the enzyme can cleave anywhere on the length of the molecule.[27-28] Thus, random cleaving from micrococcal nuclease on intact chromatin will leave a distribution of size fragments shown as distinct bands in an electrophoresis gel. These bands indicate post-digestion chromosome fragments of differing nucleosome count. If the chromosome does not contain intact nucleosomes, a smear is created as the micrococcal nuclease enzyme is no longer restricted to cleaving at any location along the molecule. ***Figure 3.4*** below shows the results we obtained

from the native M0-91 and HeLa chromatin extractions post micrococcal treatment.

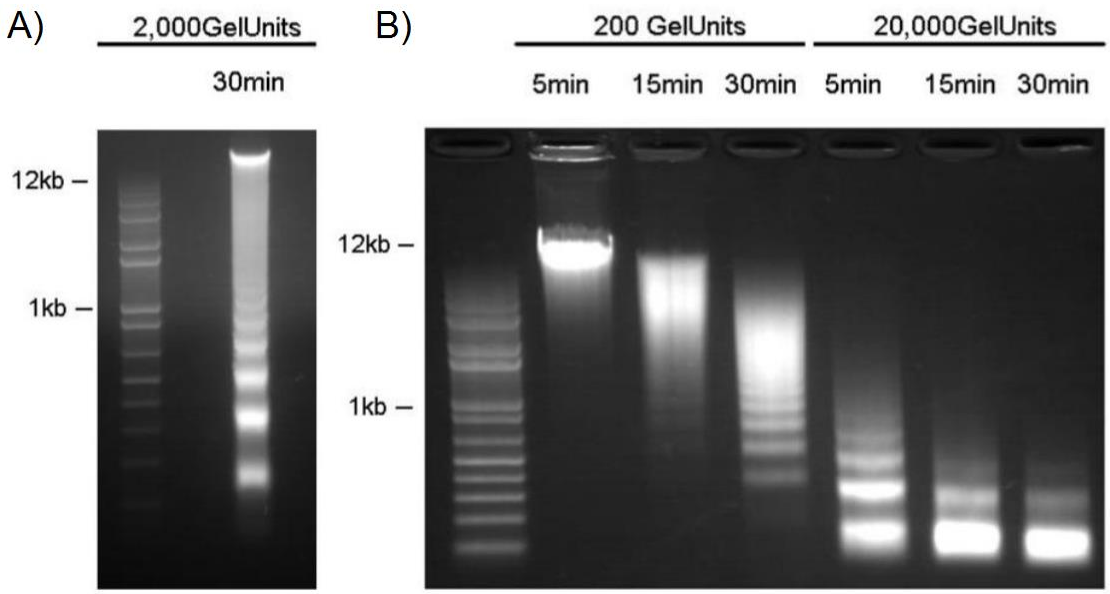

Figure 3.4 – Sizing Micrococcal Digested Chromatin on Gel

(Modified from Cerf A. et. al., 2012)

Banding patterns are visible after micrococcal digestion of native chromatin extractions for chromatin extracted from both M0-91 cells as well as HeLa-WT cells.

Combed chromatin was transfer printed onto silanized cover slips and characterized with AFM. The height of the nucleosomes on the combed native chromatin was measured to be 10.4nm ±1.8nm and the height of the DNA

backbone measured 1.76nm ±0.38nm. These results are in accordance with theoretical values of nucleosomes, 10nm, as well as bare DNA, 2nm.[29] Thus, this suggests that the chromatin combing and transfer printing process isolated single molecules rather than bundled ropes. Furthermore, the AFM images shown in ***figure 3.5*** below confirm the integrity of the nucleosomes as validated by the gel studies.

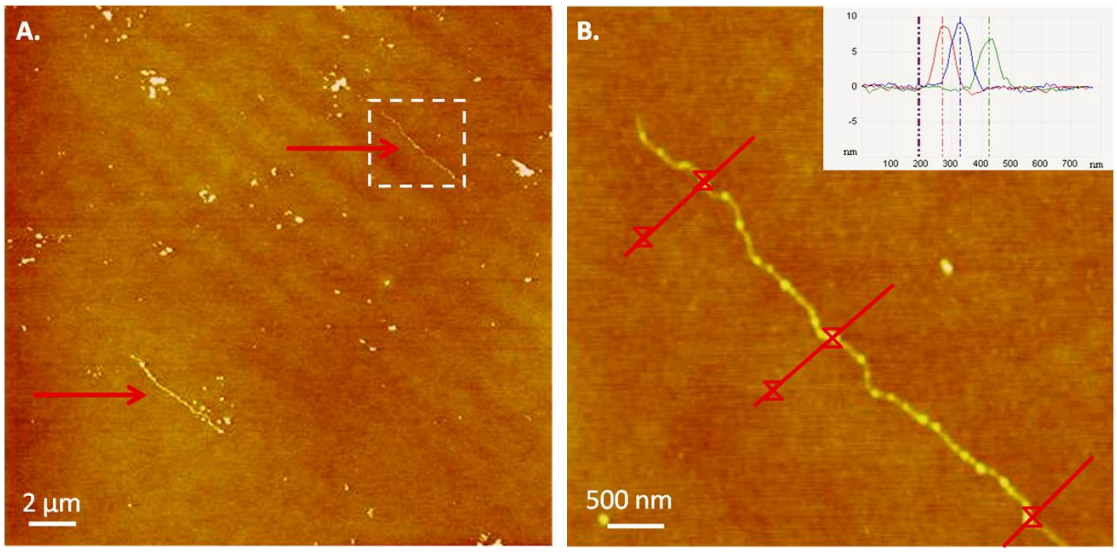

Figure 3.5 – AFM Imaging and Characterization of Combed Chromatin

(Modified from Cerf A. et. al., 2012)

(A) Chromatin arrays visible on AFM from large area scan identified by the red arrow markers. The white dotted line indicates the field of view shown in (B), where Z-axis analysis reveals that the nucleosome height and DNA backbone height correspond to

theoretical values thereby confirming the chromatin molecule observed is a single molecule.

3.3.3 Imaging Histones and Epigenetic Marks on Chromatin Arrays

Native chromatin samples pre-labeled with fluorescent antibodies bound to histone H3 were combed and imaged with a 100X oil-immersion lens on an IX-70 Olympus Inverted Microscope (Olympus). By labeling the DNA backbone with a separate color dye as to not cause spectral overlap, histone H3 can be visualized on native chromatin. This can be seen in ***figure 3.6*** below with M0-91 native chromatin that has been combed into an array. The DNA backbone imaged at 488nm is compiled with the histone H3 image taken at 650nm to form a composite image transposing histone positions on the elongated chromosomal DNA. To verify that the AlexaFluor647-tagged secondary antibody bound to the primary histone H3 antibody specifically, secondary antibody was incubated with native chromatin in the absence of the primary antibody. We found very minimal instances of non-specific secondary antibody binding to the combed chromatin despite taking no additional steps to wash away unbound or non-specifically bound secondary antibody after molecular combing or transfer printing. We then took one step further to demonstrate the

labeling of epigenetic histone modifications on combed native M0-91 chromatin with a 3-color colocalization between the DNA backbone, histone H3, and H3K9me2 as shown in ***figure 3.7***.

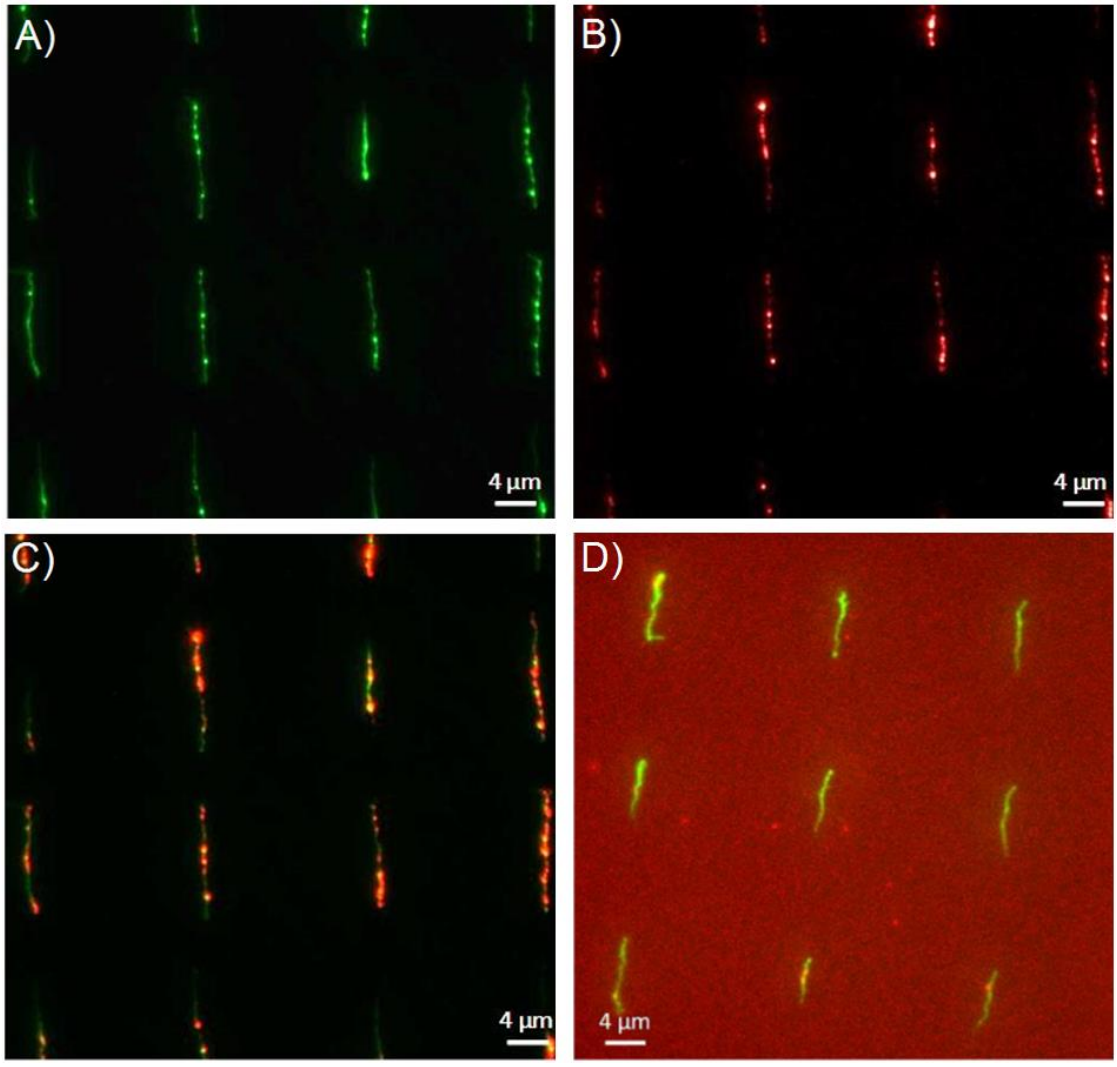

Figure 3.6 – Imaging Histone H3 on M0-91 Native Chromatin Arrays

(Modified from Cerf A. et. al., 2012)

(A) Native chromatin extracted from M0-91 cell line labeled with YOYO-1 intercalating dye allows for the fluorescence imaging of the DNA backbone. (B) The same field of view with AlexaFluor647 labeled histone H3 positions. (C) Composite images showing colocalization of histone H3 (red) on the DNA backbone (green). (D) Taken from a separate experiment where secondary antibody was incubated with M0-91 chromatin in the absence of primary antibody to show lack of non-specific binding. (Combing direction from top to bottom.)

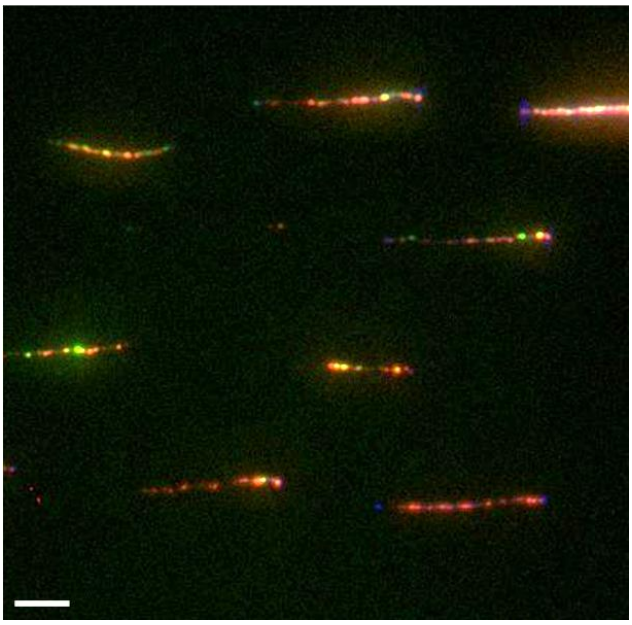

Figure 3.7 – Fluorescently Tagged Antibody Labeling of Histones and Epigenetic Modifications on Native Chromatin Arrays

Here, the chromosomal DNA backbone was labeled with DAPI dye, histone H3 was labeled with AlexaFluor647, and H3K9me2 was labeled with AlexaFluor488. All three labels can be seen in a colocalized image with combed M0-91 native chromatin. (Combing direction from right to left.)

A correlation was noted between the site of antibody tagged histones and an increased fluorescence intensity of the DNA backbone as stained with intercalating dye. We postulate this could be the result of a number of factors, the first of which is that at each labeled histone site there would be a higher spatial concentration of DNA accounted for by the DNA wrapping around the nucleosome. However, varying fluorescence intensities between these nucleosome sites may suggest a higher order structure as opposed to a perfectly linearized beads-on-a-string chromatin fragment. Possibilities include a kink or knot in the chromatin molecules combed, an observation that has been made by others on phage DNA when combing at various speeds and pH on silanized glass.[30-31] Alternatively, it is possible that although AFM results have shown our ability to comb single chromatin molecules, the fluorescence micrographs taken in ***figure 3.6*** and ***figure 3.7*** are of chromatin bundles rather than single molecules. To further elucidate these theories, it is critical to move beyond diffraction limited imaging. Coupling chromatin combing with localization microscopy techniques such as Stochastic Optical Reconstruction Microscopy

(STORM) is essential to achieve resolution scales required to accurately determine histone modification position and nucleosome distribution.

3.4 ALTERNATIVE LINEARIZAITON TECHNIQUES EXPLORED

Alternative native chromatin immobilization techniques were explored upon achieving chromatin extraction lengths of over 500kb. The motivation for extracting ultra-long beads-on-a-string chromatin was for lowering the required amount of input sample for optical mapping. Conventionally, optical mapping requires strands of DNA or chromatin to be pieced together by overlapping regions. By reducing the sum of overlapping regions necessary to reproduce the genome through increasing strand lengths, less input material is required. Among a number of linearization techniques explored including capillary flow cells and droplet drying, shear flow yielded the best results for elongation of ultra-long native chromatin. Here, a 15ul droplet of sample containing the ultra-long chromatin is micropipetted onto a 22mm x 22mm glass cover slip. Another cover slip of the same dimensions is then laid directly on top of the sample droplet thereby squeezing the contents of the droplet to span the entire area in between the top and bottom cover slips. This creates a

hydrodynamic shear force that extends molecules in solution as shown with chromatin extracted from HeLa-GFP cells in ***figure 3.8*** below.

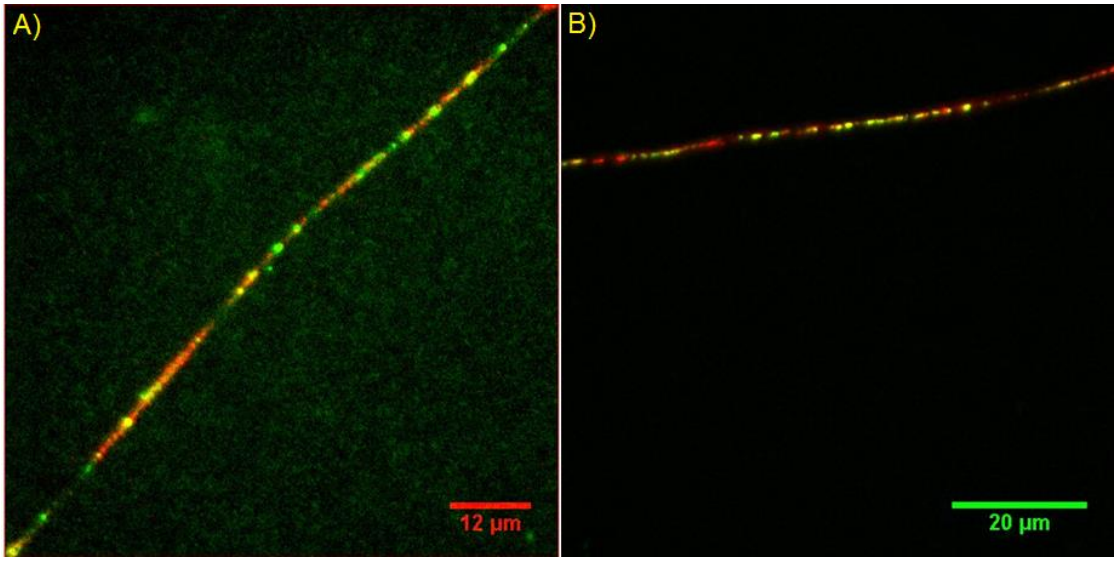

Figure 3.8 – Ultra-Long Native HeLa-GFP Chromatin Elongated via Shear Flow Between Glass Coverslips

Native chromatin extracted from HeLa-GFP cell lines expressing green fluorescent protein (green) labeled with BOBO-3 intercalating dye (red). These ultra-long fragments span a minimum of 500 kb to upwards of 1.3 Mb and are elongated via shear flow.

Unfortunately, all alternative linearization techniques explored were less reliable in comparison to patterned array molecular combing. Challenges

include severe molecule bundling as well as often an increased background fluorescence. One advantage associate with molecular combing on a hydrophobic polymer stamp was that as combing occurred, unbound dye traveled with the fluid rather than remaining on the surface. In techniques such as droplet drying and shear flow, free dye in solution was not washed out thereby causing an often high level of background fluorescence. On the other hand, flow cell techniques suffered heavily from incomplete elongation. Although these issues encountered for the alternative techniques explored could be ameliorated with extensive parameter optimization, it was ultimately decided to be too unreliable as a platform for mapping based diagnostic.

3.5 CONCLUSION

The ability to molecular comb native chromatin molecules derived from human cells holds significant implications towards potential medical diagnostics based on rapid assessment and mapping of a patient's epigenome. To this end, we have taken the first steps in applying molecular combing and transfer printing to native chromatin molecules extracted from human cancer cell lines M0-91 and HeLa. Furthermore, we have demonstrated the ability to fluorescently tag both histones as well as histone modifications for high resolution imaging. We

show the ability to colocalize multiple fluorescent tags on single chromatin molecules however, moving past diffraction-limited optics is necessary to improving mapping accuracy. Still, challenges remain in correlating chromatin molecule lengths as measured with microscopy to a base pair count. This is because unlike bare DNA, the DNA in chromatin is wrapped around irregularly spaced nucleosomes. Furthermore, as the DNA naturally only wrapped 1.65 turns around a nucleosome rather than one or two full turns, predicting the degree of unraveling caused by fully extending the molecule for each nucleosome is a significant challenge. Together, these challenges represent some of the shortcomings of an optical based mapping as applied to chromatin. Nonetheless, we have shown the first steps towards the overall greater effort of enabling rapid epigenome mapping based on direct visual interrogation of native mammalian chromatin.

3.6 REFERENCES

1. P. R. Langer-Safer, M. Levine, D. C. Ward, PNAS, 1982, 79, 4381-4385

2. M. R. Speicher, S. G. Ballard, D. C. Ward, Nature Genetics, 1996, 12, 368-375
3. A. Weise, P. Harbarth, U. Claussen, T. Leihr, Journal of Histochemistry, 2003, 51(4), 549-551
4. H. Labit, A. Goldar, G. Guilbaud, C. Douarche, O. Hyrien, K. Marheineke, Biotechniques, 2008, 45(6), 649-658
5. K. Klinger, G. Landes, D. Shook, R. Harvey, L. Lopez, P. Locke, T. Lerner, R. Osathanondh, B. Leverone, T. Houseal, K. Pavelka, W. Dackowski, Am. J. Hum. Genet., 1992, 51, 55-62
6. N. Moosani, H. A. Pattison, M. D. Carter, D. M. Cox, A. W. Rademaker, R. H. Martin, Fertility and Sterility, 64(4), 811-817
7. D. Pinkel, J. Landegent, C. Collins, J. Fuscoe, R. Segraves, J. Lucas, J. Gray, PNAS, 1988, 85, 9138-9142
8. P. S. Langedijk, F. Schut, G. J. Jansen, G. C. Raangs, G. R. Kamphuis, M. H. F. Wilkinson, G. W. Welling, App. and Env. Microbiol., 1995, 61(8), 3069-3075
9. D. J. Wolff, A. Bagg, L. D. Cooley, G. W. Dewald, B. A. Hirsch, P. B. Jacky, K. W. Rao, P. N. Rao, J. of Mol. Diag., 2007, 9(2), 134-143
10. J. Drach, J. Schuster, H. Nowotny, J. Angerler, F. Rosenthal, M. Fiegl, C. Rothermundt, A. Gsur, U. Jager, R. Heinz, K. Lechner, H. Ludwig, H. Huber, Cancer Research, 1995, 55(17), 3854-3859

11. R. S. Cajulis, G. K. Haines, D. F. Hidvegi, K. McVary, J. W. Bacus, Diagnostic Cytopathology, 1995, 13(3), 214-223
12. D. S. Rupa, L. Hasegawa, D. A. Eastmond, Cancer Research, 1995, 55(3), 640-645
13. H. H. Q. Heng, J. Squire, L. C. Tsui, PNAS, 1992, 89, 9509-9513
14. Y. S. Fan, L. M. Davis, T. B. Shows, PNAS, 1990, 87, 6223-6227
15. I. Williamson, S. Berlivet, R. Eskeland, S. Boyle, R. S. Illingworth, D. Paquette, J. Dostie, W. A. Bickmore, Genes and Development, 2014, 28, 2778-2791
16. Y. Huan, T. Zhang, P. Thammapichai, Y. Weng, J. Jiang, Genetics, 2015, 200, 771-779
17. R. J. Florijn, L. A. Bonden, H. Vrolijk, J. Wiegant, J. W. Vaandrager, F. Baas, J. T. den Dunnen, H. J. Tanker, G. J. van Ommen, A. K. Raap, Hum. Mol. Genet., 1995, 4(5), 831-836
18. S. S. P. Rao, M. H. Huntley, N. C. Durand, E. K. Stamenova, I. D. Bochkov, J. T. Robinson, A. L. Sanborn, I. Machol, A. D. Omer, E. S. Lander, E. L. Aiden, Cell, 2014, 159, 1665-1680
19. J. Guan, L. J. Lee, PNAS, 2005, 102(51), 18321-18325
20. S. Caburet, C. Conti, A. Bensimon, Biotechnology, 2002, 20(8), 344-350
21. L. Jiang, H. Dong, W. Hu, Soft Matter, 2011, 7, 1615-1630

22. A. Cerf, B. R. Cipriany, J. J. Benitez, H. G. Craighead, Anal. Chem., 2011, 83, 8073-8077

23. X. Michalet, R. Ekong, F. Fougerousse, S. Rousseaux, C. Schurra, N. Hornigold, M. van Slegtenhorst, J. Wolfe, S. Povey, J. S. Beckmann, A. Bensimon, Science, 1997, 277, 1518-1523

24. K. Monier, M. Hazzouri, R. Ekong, S. Rousseaux, Technical Tips, 1998, 3, 75-79

25. K. Monier, X. Michalet, J. Lamartine, C. Schurra, F. Heitzmann, L. Yin, R. Cinti, B. S. Sylla, M. Creaven, G. Porta, C. Vourc'h, M. R. Nicoud, A. Bensimon, G. Romeo, Cytogenet. Cell Genet., 1998, 81, 259-264

26. S. B. Smith, Y. Cui, C. Bustamante, Science, 1996, 271(5250), 795-799

27. D. J. Clark, J. Biomol. Struct. Dyn., 2010, 27(6), 781-793

28. D. Z. Staynov, Nucleic Acids Research, 2000, 28(16), 3092-3099

29. K. Maeshima, R. Imai, S. Tamura, T. Nozaki, Chromsoma, 2014, 123(3), 225-237

30. A. Benke, M. Mertig, W. Pompe, Nanotechnology, 2011, 22, 035304 (8pp), 1-8

31. A. Bensimon, A. Simon, A. Chiffaudel, V. Croquette, F. Helsot, D. Bensimon, Science, 1994, 265, 2096-2098

32. A. Cerf, H. C. Tian, H. G. Craighead, ACS Nano, 2012, 6 (9), 7928–7934

CHAPTER 4

MICROFLUIDIC DEVICE FOR CELL PROCESSING, DNA EXTRACTION, AND IN-CHANNEL IMAGING

4.1 INTRODUCTION

While direct approaches to genomic and epigenomic mapping covered in previous chapters are being explored as alternatives to well-established techniques, they require a large amount of cellular input. As a result, these direct mapping approaches become unsuitable when considering diagnostic applications with limited starting material. One such application is genetic and epigenetic analysis of circulating tumor cells (CTCs). CTCs are highly malignant cancer cells known for their ability to translocate through the endothelial linings of blood vessel walls and their ability to survive undetected in the blood stream.[1-3] By these means, CTCs from one tumor site can migrate throughout the vasculature and exit to seed satellite tumors in separate organs or tissues. This process, known as metastasis, is responsible for over 90% of cancer related deaths in humans.[4] Thus, understanding the unique properties

of a patient's CTCs, such as the genetic and epigenetic profile, may lead to improved cancer diagnostics and treatments.[5]

However, reliably isolating CTCs remains challenging due to their ultra-low frequency. Found in our bodies at concentrations of only 1 to 10 cells per milliliter of whole blood, analyzing CTCs becomes impractical with many existing technologies that require thousands or even hundreds of cells as input.[6] To this end, numerous microfabricated technologies have been developed for single cell capture and isolation of nucleic acids as preparation for analysis platforms such as sequencing.[7-12] Of the microfabricated technologies that are capable of single cell DNA extraction and purification from few cells, most do so through solid phase extraction (SPE).

In SPE based techniques, DNA is purified by selective binding to functionalized surfaces, such as that of microbeads, and then washing out the cell lysate. Although these mechanisms can be easily integrated with a wide range of single cell isolation techniques such as fluorescence activated cell sorting (FACS), the process ultimately relies on the binding interaction between the solid phase matrices and the target DNA. Despite a number of recent advances made to improve the binding affinity between the DNA and matrix such as by adjusting buffer composition, pH, and temperature, it is still a fundamentally equilibrium

driven interaction where DNA loss is likely.[13-15] Furthermore, binding between the solid phase matrices and DNA can be subject to interference by proteins and cellular debris contained within the lysate. Consequently, SPE based methods have only achieved DNA extraction efficiencies of between 50% and 80%.[16-19] Although this recovery efficiency may be sufficient for genomic analysis of cell populations, the limited gene copies available within a single cell requires a higher precision approach.

In this chapter, we describe a single integrated microfluidic device capable of single cell and small cell population capture, DNA extraction, purification, in-channel labeling, in-channel imaging, and off-chip collection. The device presented is a simple two port, valveless, PDMS microfluidic chip that operates based on using a micropillar array to achieve the aforementioned capabilities. Although microfabricated obstacle arrays have been extensively studied as applied to DNA size filtration and linearize, cellular genomic DNA extraction using micropillar array has yet to be demonstrated. Unlike SPE-based methods, micropillar array DNA extraction is a purely physical process that does not require any DNA binding or surface functionalization. Here, we use the micropillar array based microfluidic device to extract and purify DNA from varying cell counts of human cancer cell lines. We assess off-chip collection of the extracted DNA and measure extraction efficiency through fluorescence

quantification and qPCR. Finally, we also test in-channel epigenetic modification labeling and imaging.

4.2 MATERIALS AND METHODS

4.2.1 Microfluidic Device Fabrication

Illustrated by the workflow in ***figure 4.1***, silicon molds for polydimethylsiloxane (PDMS) microfluidic devices were fabricated using standard photolithography techniques. Briefly, wafers (Ultrasil; Hayward, CA) were spin coated with Microposit S1813 photoresist (Shipley; Marlborough, MA). Device pattern was transferred onto photoresist layer using UV contact lithography (ABM contact aligner, ABM-USA; San Jose, CA). Exposed substrates were developed in 726MIF developer (Microchemicals). The microfluidic pattern was transferred onto the top silicon layer by Bosch process in a Unaxis SLR 770 deep reactive ion etching system (Unaxis USA Inc.; St. Petersburg, FL). Etch depth was determined to be 20 – 25 μm using a P10 profilometer (KLA Tencor; Milipitas, CA) and a Zygo otical profilometer (Zygo Corporation; Middlefield, CT). A monolayer of (1H,1H,2H,2H-Perfluorooctyl) Trichlorosilane was deposited on the etched wafers in a MVD100 molecular wafer deposition system (Applied Microstructures; San Jose, CA) to prevent adhesion of PDMS to the mold.

Sylgard 184 (Dow Corning; Midland, MI) PDMS base resin was mixed with the curing agent at a 10:1 ratio, degassed under vacuum at room temperature, poured onto the master until a 1cm thick layer was achieved, and then heat cured for 45 minutes at 150C. The elastomer casting was then peeled off the mold and access holes to the input and outputs of the microchannels were created via 1.5mm biopsy punch (Sklar Instruments; West Chester, PA). To complete channel fabrication, the patterned PDMS was treated with oxygen plasma for 5 minute and bonded to a 500 μm thick fused silica wafer (Mark Optics; Santa Ana, CA).

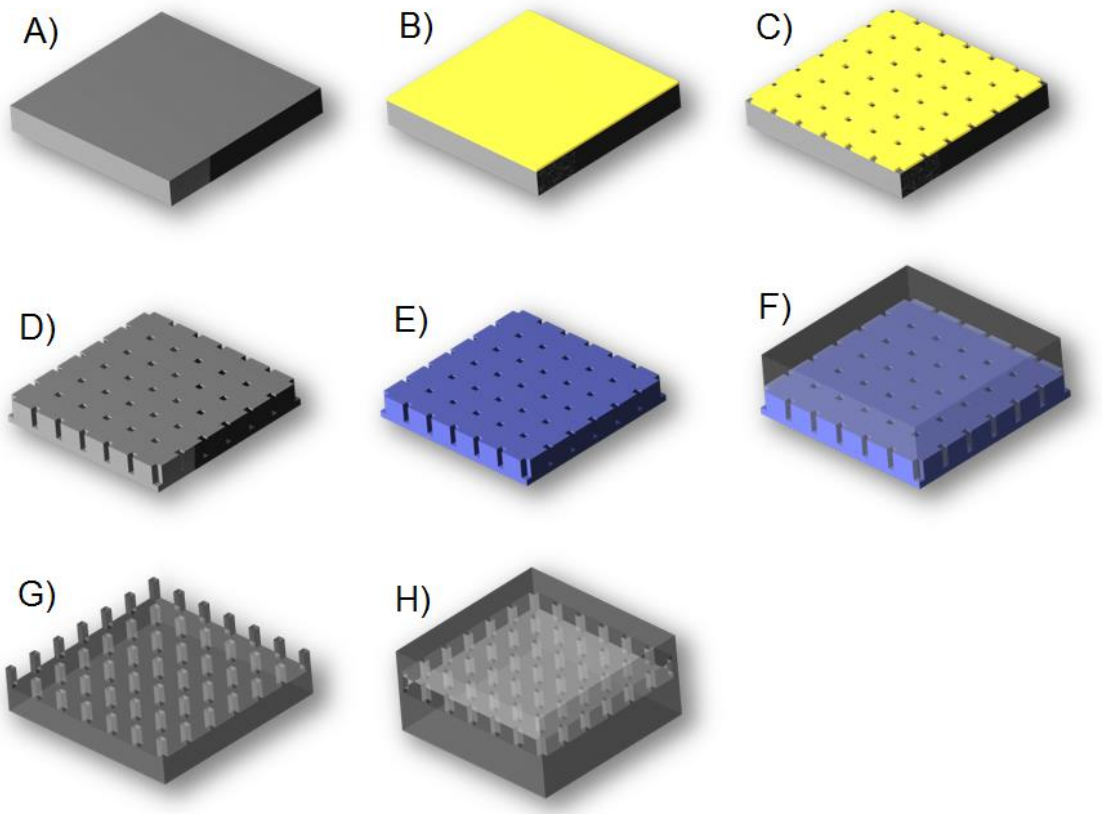

Figure 4.1 – Microfluidic Device Photolithography Process Outline

Starting with a clean 4-inch diameter silicon wafer surface (A) photoresist (yellow) is spin coated onto the surface (B). The pattern is then exposed and developed (C) before the photoresist is stripped and the wafer is the wafer is etched (D). To prevent polymer binding on the surface during mold-casting, a thin layer of FOTS is vapor deposited onto the silicon master (E). To fabricate the polymer microfluidic device, PDMS base mixed with curing agent is poured onto the silicon master (F). After removing the cured silicon slab containing the pattern (G), it is then plasma bonded onto a glass silica wafer to create the final microfluidic device (H).

4.2.2 Cell Culture

Myeloid leukemia cell line M0-91 and cervical cancer cell line HeLa wild type cells were cultured in the same conditions outlined in ***section 3.2.2*** of ***Chapter 3***. Prior to introduction into the channel, M0-91 cell lines were taken directly from an 80%-90% confluency culture flask containing media and diluted 1:50 into 1x PBS buffer. Similarly, for HeLa-WT cells post trypsinizing, cell pellets were resuspended in 5ml of 1x PBS buffer and further diluted at 1:50 in 1x PBS buffer prior. To help prevent cell adherence onto the inner walls of the microfluidic channels, the devices were primed with 1% BSA at room temperature for 2 hours prior to the addition of cells.

4.2.3 Microfluidic device operation

The microfluidic device was first treated with 5 minutes of oxygen plasma using a Harrick Plasma Cleaner to facilitate initial channel priming. 20ul of 100% ethanol was used to prime the device immediately after oxygen plasma cleaning. The device was then secured to the stage of an Olympus IX71 microscope via masking tape with channels above the objective. Hard plastic

PEEK 1/16th inch OD Tubing (Upchurch Scientific) was connected to an empty Hamilton 81110 glass 0.25ml syringe (Hamilton) via plastic flange fittings (JL White). 100% H_2O was manually pulled into the syringe through the PEEK tubing until there was approximately 50ul within the syringe. The syringe was mounted on a PHD Ultra Pump (Harvard Apparatus) and secured via the pump's inbuilt syringe clamps. The water was infused out of the PEEK tubing until a 10ul droplet formed at the output before inserting the tubing into the output port of the microfluidic device. In this step, extra care was taken to ensure a droplet-to-droplet connection was made as to not introduce any air bubbles into the microfluidic system. After tube insertion, the pump was set to withdraw mode at 60nl min^{-1} and water was flown into the microfluidic device for 5 minutes. The input port was then loaded with priming buffer consisting of 1x PBS and 1% bovine serum albumin (BSA) for 2 hours to block the inner channel walls from nonspecific protein adhesion of the cells. Next, channel was flushed with 1x PBS buffer for 5 minutes. Cells were loaded into the device again via the input port. Images of captured cells were taken at this time along with inspecting the length of the device to ensure no randomly adhered cells were lodged either upstream or downstream of the cell capture region. Lysis buffer consisting of 1% (wt) sodium dodecyl sulfate (SDS) surfactant in PBS was flowed into the channel at a constant speed of 15nl min^{-1} for a total of 5 minutes to lyse the cells. In contrast to DNA extraction, chromatin extraction was

performed with a different lysis buffer, containing 1% (vl) Triton-X in 1x PBS rather than using any SDS. In either case, the lysis buffer was then washed out of the channels by flowing water for into the device 10 minutes. The device and the entrapped DNA or chromatin is now ready for imaging or off-chip collection following the procedures described in the following sections.

4.2.4 In-Channel Imaging of DNA and Chromatin

DNA was imaged with a number of different intercalating dyes. To image the DNA with SYBR-Gold intercalating dye (Invitrogen), 1:10,000 dilution of SYBR-Gold dye was prepared in 1x TE buffer at room temperature and flowed into the channel for a minimum of 20 minutes in the absence of light prior to fluorescence imaging. Similarly, SYBR-Green intercalating dye (Invitrogen) was also used in various experiments to stain DNA following the same protocols as SYBR-Gold dye. Alternatively, Quant-iT PicoGreen intercalating dye (Thermo Fisher Scientific) has also been used to stain bare DNA by diluting the dye at 1:200 in 1x TE buffer. PicoGreen dye was flowed into the channel for a minimum of 15 minutes in the absence of light prior to imaging.

Histone H3 labeling of in-channel chromatin was done with Histone H3 (Abcam) fluorescently conjugated with Alexa Fluor 647 using a protein labeling kit (Thermo Fisher Scientific). The antibody was flowed into the channel at a concentration of 1.5ug ml^{-1} in priming buffer supplemented with 0.05% Tween-20 and incubated for one hour. Excess and unbound antibody was washed out of the channel prior to imaging. Similarly, H3K9me2 was labeled with Alexa Fluor 647 conjugated H3K9me2 Antibody (Abcam) and incubated for an hour before imaging as well. A concentration of 10 ug ml^{-1} was used for H3K9me2 in the priming buffer with 0.05% Tween-20. To calibrate the fluorescence intensity for fluorescence based quantification of H3 or H3K9me2, Cy5 dye was imaged at 30nM and 30μM concentrations in blank channels free of cells and biological materials.

4.2.5 Off-Chip DNA Collection and Quantification

In order to extract the DNA from its entangled state within the microfluidic device, it must be enzymatically cleaved and released downstream. 1ul of BAM-HI enzyme (New England BioLabs) was diluted in 100ul of Buffer 2 (New England BioLabs) and preheated to 37C in a water bath. The enzyme was then loaded into the input port of the microfluidic device and flowed at a rate of

100nl min^{-1}. To keep the reaction as close to 37C as possible, the microfluidic device was heated via heat gun (Master). Fragmented DNA was collected into the 0.25ml syringe and then transferred into a 1.5ml eppendorf tube. DNA was fluorescently stained with PicoGreen dye at a 1:200 dilution and then quantified via a NanoDrop 3300 Fluorospectrometer (NanoDrop). Calibration curves for the fluorospectrometer were generated using identical staining protocols with a series dilution of T4 bacteriophage DNA (100ng ml^{-1}, 50ng ml^{-1}, 25ng ml^{-1}, 12.5ng ml^{-1}, 6.25ng ml^{-1}, and 3.13ng ml^{-1}). Measures were blanked with a background set by 1:200 dilution of PicoGreen in water.

For experiments pertaining to the qPCR, M0-91 gDNA standards were obtained from phenol chloroform bulk-extraction of 2 x 10^6 M0-91 cells. The DNA was further purified via ethanol precipitation and stored in 1x TE buffer at 4C until use. To create qPCR DNA standards, the bulk-extracted M0-91 gDNA was diluted to samples with a DNA mass of 3.3pg, 6.6pg, and 9.9pg. DNA extracted from single M0-91 cells on chip was released via enzymatic cleaving with Hind-III at 37C for 1 hour. Hind-III was then heat inactivated by submerging the collected sample in 65C water bath for 10 minutes. Whole genome amplification was completed with the Single Cell WGA kit (New England BioLabs). qPCR was performed with a LightCycler 480 Real-Time PCR thermocycler (Roche).

4.3 RESULTS

4.3.1 Micropillar Array Channel Design

The micropillar array in the device served dual purposes of both size-based cell capture as well as entrapment of DNA and chromatin upon cell lysis by means of physical entanglement. Shown in ***figure 4.2*** below, various dimensions and designs of the micropillar region were explored including progressively higher density pillars, random pillar placement, and changing the channels widths to accommodate different cell counts.

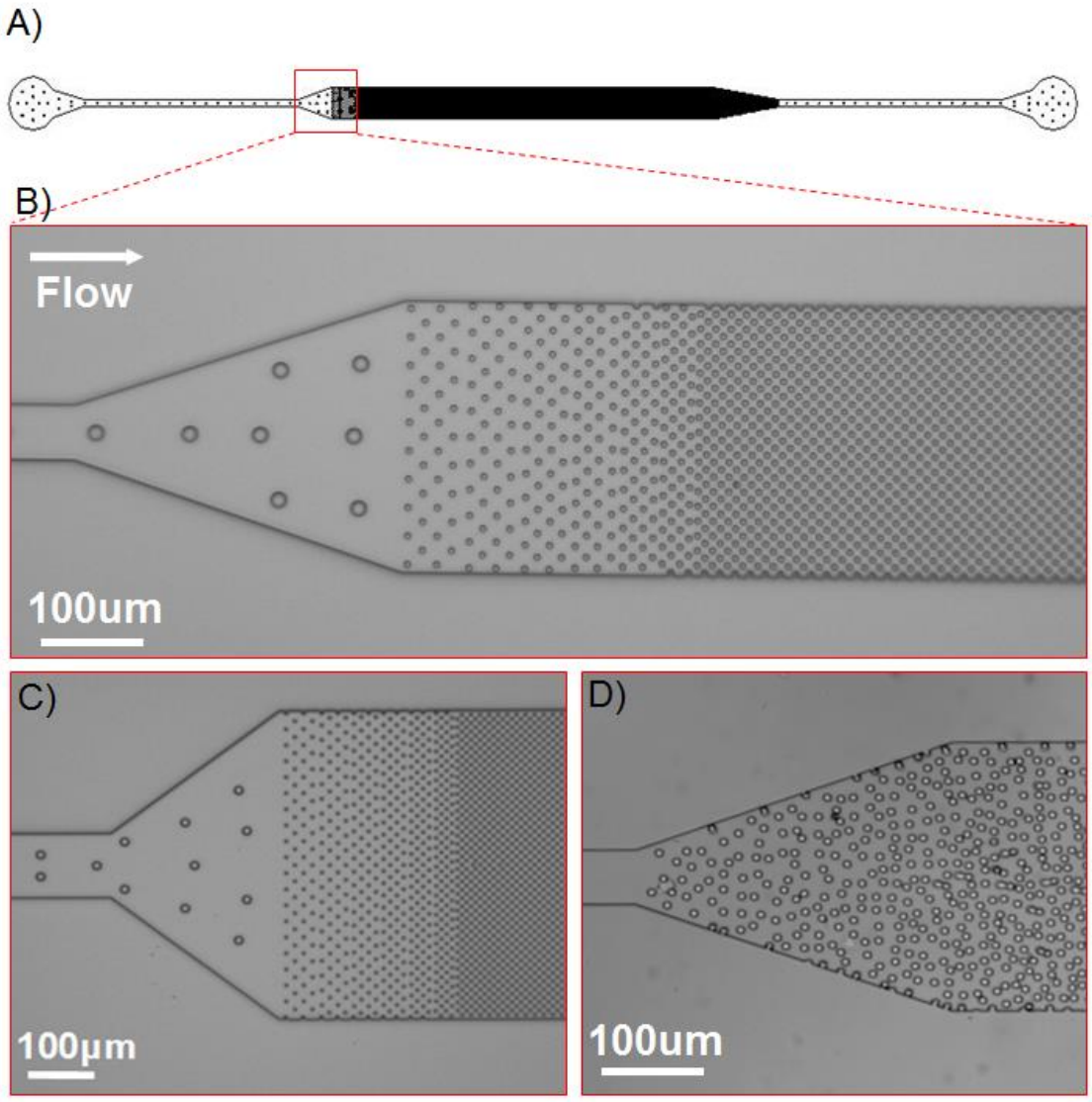

Figure 4.2 – Cell Capture and DNA Extraction Region Designs

(A) Zoomed out crop of the microfluidic channel design consisting of an input port (left), output port (right), and tapered narrow channels leading into and out of the micropillar array region where cell capture and DNA extraction takes place. (B) A 250µm wide channel with progressively denser spaced (15µm spacing down to 2µm spacing) micropillars. (B) A 500µm wide channel made to accommodate larger cell

counts. (C) Randomized pillar arrangement design as an alternative to the ordered arrays design was tested for its ability to increase DNA capture efficiency.

Support pillars 15µm in diameter were spaced sparsely upstream of micropillar arrays to prevent channel collapse from the negative pressure generated by the withdraw of the syringe pump. Smaller micropillars spanned 5µm in diameter and at the highest density configuration, were spaced only 2µm apart from one another. The channel leading into and out of the device was tapered to a narrower width canals to increase flow rate of the cells from the input port to the micropillar capture region while reducing collision velocity once the cells entered into the broader regions of the channel where the pillar array was located. However, we later found that the support structures in the constricted narrower channel regions leading into and out of the micropillar array region often caused undesired cell adhesion, collision, and DNA entanglement. In all of the micropillar configurations tested across devices, the arrangement of 5µm diameter microposts with respect to one another held no visible impact on the DNA capture capabilities.

Various designs for the regions immediately downstream adjacent to the cell capture and DNA extraction micropillar arrays were also explored as shown in ***figure 4.3***. The reason for doing so is that our original design consisted of

extending the tightly packed pillar array all the way until the tapered output canals but we found that for longer channels, extended lengths of highly dense micropillar arrays began to impede the hydrodynamic flow rates within the channel. Alternative designs were explored with the effort to have the pillars downstream of the cell capture and DNA extraction region act as both support structures to prevent channel collapse as well as points of electrostatic contact for strands of DNA that are tethered upstream. However, we observed no change to the tethering behavior or extraction capabilities of the DNA extraction region by removing these downstream structures in later iterations of the microfluidic device.

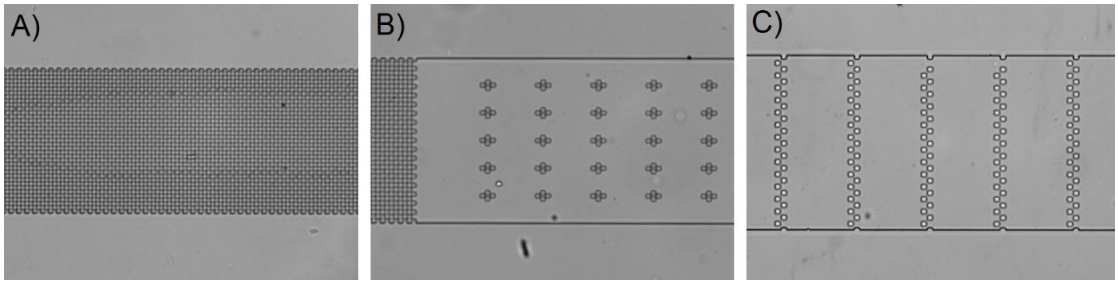

Figure 4.3 – Downstream Micropillar Configuration Designs

(A) Dense micropillar arrays extended fully downstream impedes flow rate at longer channel lengths. (B) Alternative design with clusters of 4 micropillars spaced in a diamond aligned in rows of 5 periodically spaced downstream of the main cell capture

and DNA extraction regions. (C) Staggered pillars aligned in periodic rows to help provide additional points of electrostatic contact for DNA molecules suspended upstream.

When mold-casting the PDMS slab to create the microfluidic device, we observed that the speed at which the cured PDMS was peeled from the silicon master had the potential to affect the final channels. Peeling the PDMS from the device too quickly was detrimental in that many micropillars would be ripped from the PDMS slab and remain lodged within the features of the silicon master, as shown by ***figure 4.4***, thereby rendering that silicon master unfit for further mold-casting. This issue was alleviated by vapor depositing a monolayer of fluoroalkylsilane FOTS on the silicon master after photolithography.

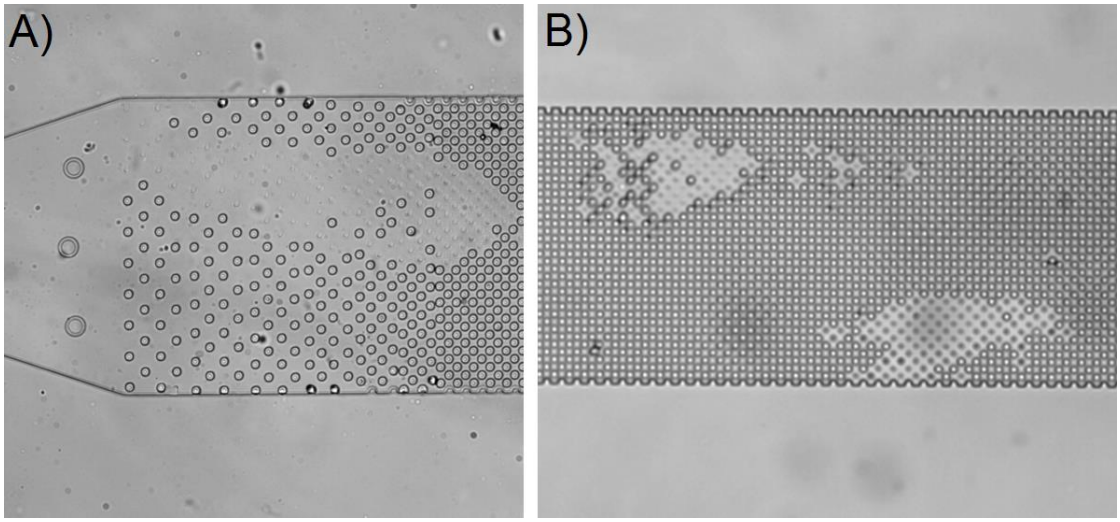

Figure 4.4 – Micropillar Stripping in Failed Devices

Peeling the PDMS slab from the silicon master often causes micropillars to be stripped and lodged within the silicon mold. Silicon masters were coated with FOTS to help prevent this occurrence.

Lastly, because smaller features etch slower than larger features in photolithography, taller channel designs of heights exceeding 25μm often exhibited micropillars that were unable to reach the floor of the device as depicted by ***figure 4.5***. This produced the effect of cells being able to deform and migrate through the PDMS pillars unimpeded at higher flow rates. A combination of lower channel heights, softer PDMS made from higher base-to-curing agent mixing ratios, and slower flow rates of less than 3ul min-1 were used to combat this problem. Additionally, pressure was applied to clamp the PDMS slab to the glass after plasma bonding to further aid the pillars in making contact with the glass surface.

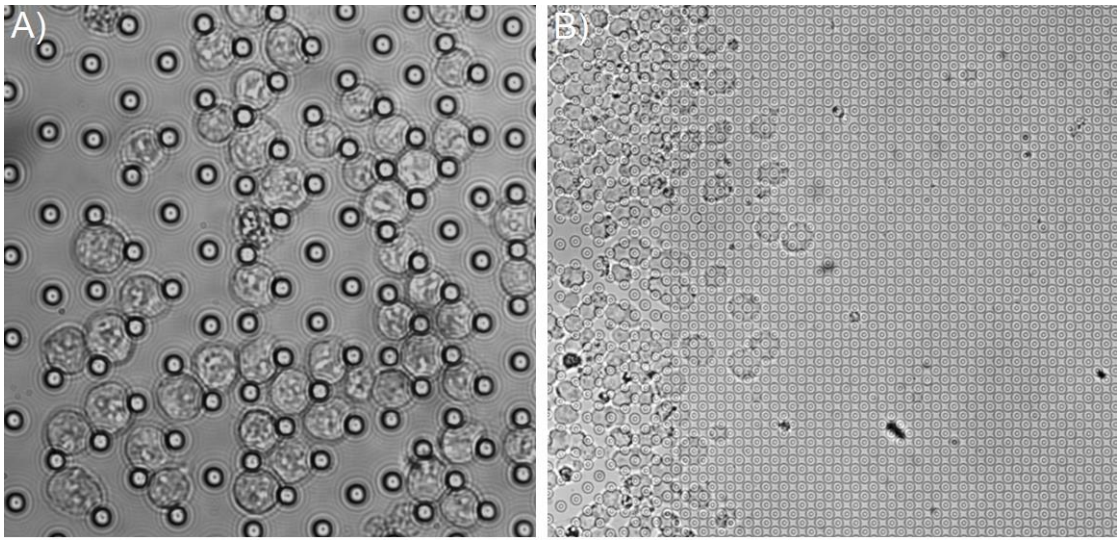

Figure 4.5 – Cells Squeezing Between the Micropillars and Glass on Devices with Heights Exceeding 25µm

(A) M0-91 cells trapped in the micropillar array on a device with micropillars that reach all the way to the glass. (B) Aspect ratio dependent etching rate of silicon mold resulting in micropillars that were too short thus allowing M0-91 cells to pass under them resulting in failed cell capture. (In both panels, fluidic flow moves from the direction of left to right).

4.3.2 DNA Extraction and Staining

The unique novelty of our microfluidic device is DNA capture upon cell lysis through physical entanglement. This process is done in complete absence of any

chemical modifications to the cell, DNA, or to the inner channel walls of the device. Relying solely on the size of the DNA molecules, the micropillar array entangles the DNA much similar to how spaghetti may become entangled on a fork. As the surfactant within the lysis buffer denatures the histones packing the DNA into higher order structures, the DNA unravels in solution as it collides into downstream pillars causing the tethering. This tethering can occur in a number of different configurations as shown in ***figure 4.6*** below. Due to the ratio between gDNA length and our micropillar diameters, the most common configuration for the tethered DNA likely resembles a molecule that is tethered at multiple locations and folding on itself.

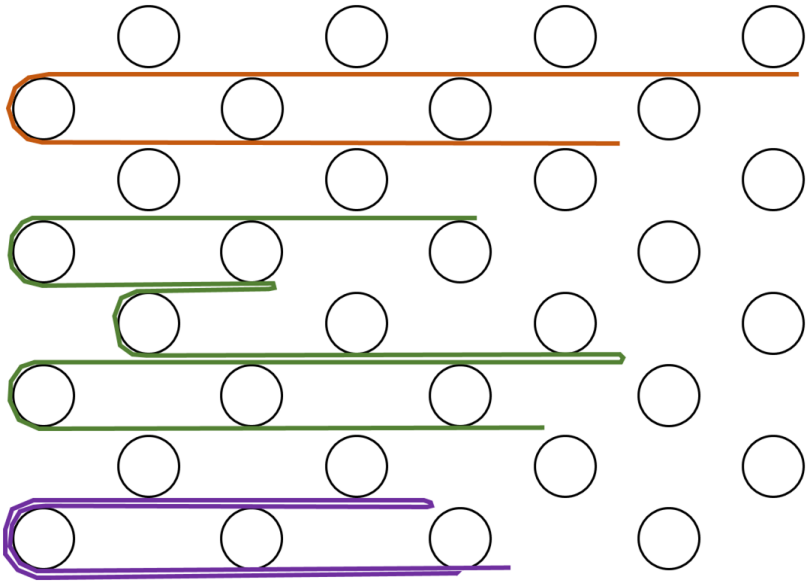

Figure 4.6 – Illustration of gDNA Tethering Configurations within Micropillar Array Post Cell Lysis

This figure represents DNA tethered on the micropillars post cell lysis in a micropillar array device where flow is moving in the direction from left to right. The simplest tethering configuration (orange) would be a molecule that is tethered at one point by a single micropillar. A variation of this configuration is if a molecule is folded on itself either once or a multiple of times yet is still tethered by a single micropillar (purple). The most likely scenario given the scale of the DNA molecules is that there are multiple levels of folding events and multiple micropillars upon which a single DNA molecule can be arrested.

Furthermore, smaller molecules such as RNA, lipids, proteins, and other cellular debris is washed away downstream leaving only the genomic DNA (gDNA) tethered on the micropillars. By this process, the extracted gDNA is also purified without the need for downstream or off-chip filtration. Furthermore, the local environment of the DNA can be controlled within our microfluidic system by simply exchanging the input fluid into our device. The tethered DNA can then be stained with fluorescent intercalating dyes such as SYBR dyes or PicoGreen dye. As shown in ***figure 4.7***, DNA from each individual cell trapped within the pillar array becomes entangled just downstream of the original position of the cell prior to the introduction of lysis buffer. Depending

on cell count, the amount of extracted DNA tethered to the pillars can be either visualized as individually resolvable strands or a meshed web.

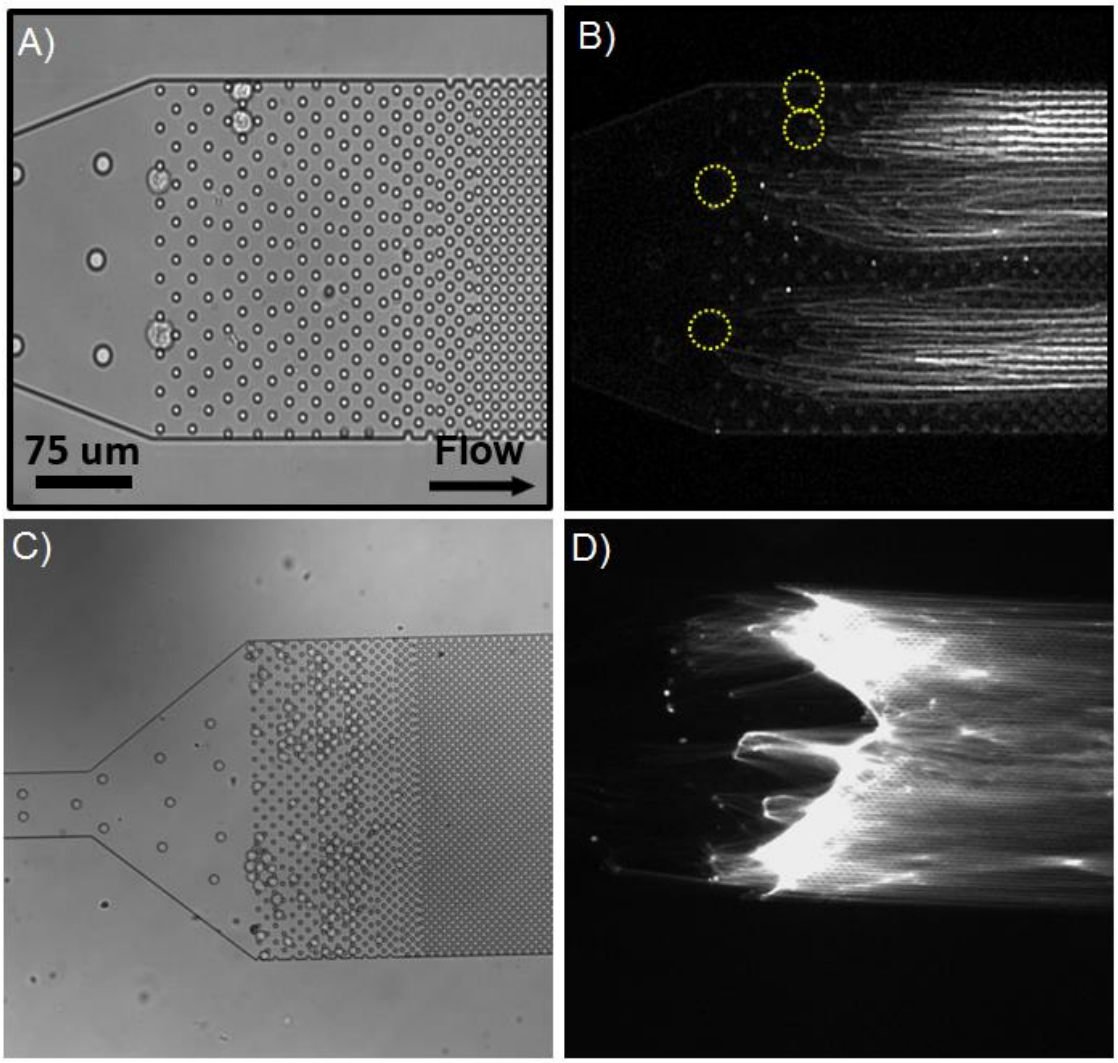

Figure 4.7 – DNA Extraction from 4 and 100+ Cells Using 250µm and 500µm Width Microfluidic Channels

(A) Four M0-91 cells are trapped within the less spatially dense cell capture region of the micropillar array. (B) The DNA from the four cells shown (positions denoted by yellow dotted line circles) in (A) are tethered on the pillars downstream after cell lysis and are visible through PicoGreen staining and fluorescence imaging. Individual DNA strands are visibly entangled around microposts. (C) 100+ M0-91 cells trapped in the cell capture region of a larger microchannel. (D) Web of DNA taken extracted from the cells shown in (C) fill the channel downstream of the cell capture region. (In all panels, fluidic flow moves from the direction of left to right).

When loading cells into the device at concentrations of above 10^4 cells ml^{-1} we encountered many cases of cells becoming adhered onto the upstream support pillars. As seen in ***figure 4.8*** below, these support pillars often trapped cells prior to the micropillar array and acted to entangle the DNA from those upstream cells during lysis. One possible issue with such events is that the lack of pillar density around the support structures may cause significant loss of gDNA during lysis, thereby inaccurately skewing cell count vs trapped DNA. Such was our reasoning for later removing upstream support structures and opting for narrower channels when redesigning our microfluidic channels for single cell experiments.

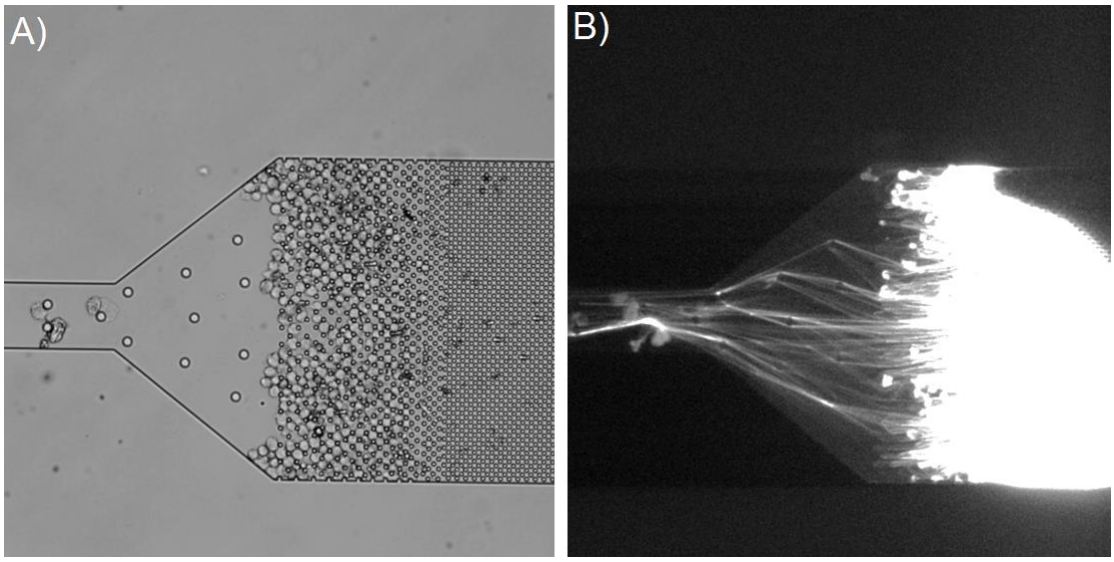

Figure 4.8 – Unintended Cell Trapping and DNA Tethering by Upstream Support Structures

(A) Cells can become adhered or trapped on upstream support structures prior to entering the micropillar array regions which then causes (B) DNA tethering on support structures upon cell lysis. (In both panels, fluidic flow moves from the direction of left to right).

4.3.3 Off-Chip DNA Fluorescence Quantification

To ensure that we were not losing significant quantities of DNA through the micropillar array DNA extraction process, we used enzymatic digestion of the

Bam-HI restriction endonuclease to cleave and release the tethered gDNA, collected it off-chip, and quantified the yield through fluorospectrometer. We first created a calibration curve for our NanoDrop 3300 fluorospectrometer through a series dilution of bacteriophage T4 DNA at 100ng, 50ng, 25ng, 12.5ng, 6.25ng, and 3.125ng ml^{-1} as shown in ***figure 4.9*** below.

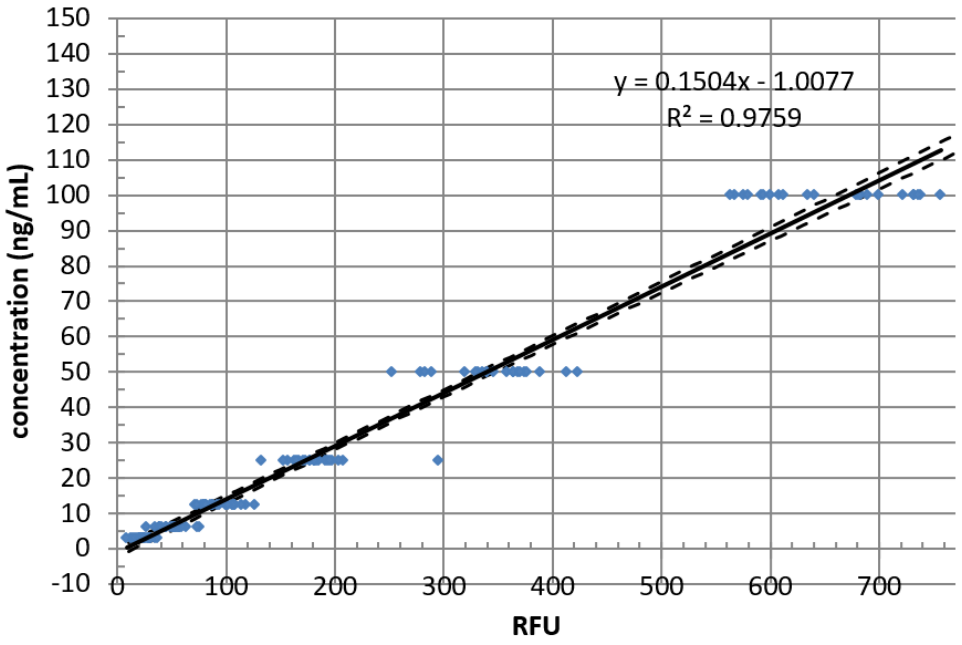

Figure 4.9 – T4 DNA Based Fluorospectrometer Calibration

N=23 for each DNA concentration in the series dilution of T4 DNA. Data taken on a NanoDrop 3300 fluorospectrometer with 2.5ul droplets for each measurement. Baseline fluorescence calibrated with water (n=25).

To collect the extracted gDNA off-chip, restriction endonuclease Bam-HI was flowed into the device to cleave and release the tethered DNA. Due to the high concentrations of enzyme used, DNA cleaving and release occurred rapidly reaching completion in under 10 minutes. This process was observable in real-time by fluorescently staining the tethered DNA with intercalating dye and imaging the micropillar arrays during the cleaving process. In ***figure 4.10*** below, gDNA extracted from M0-91 cells that have been fluorescently stained with PicoGreen intercalating dye. Within two minutes, all of the DNA originally tethered on the micropillars during the extraction process has been cleaved and released downstream.

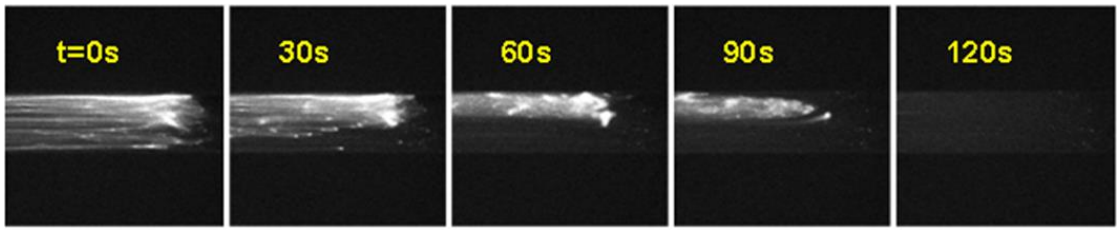

Figure 4.10 – DNA Release from Micropillars via Enzymatic Digestion

Fluorescently stained DNA extracted from 66 M0-91 cells are cleaved with restriction endonuclease Bam-HI and released downstream to be collected and quantified off-chip. The area being observed is the microfluidic pillar array region containing the tethered

gDNA. Time stamps of the frame taken from a real-time video are labeled on the corresponding panel. (In all panels, fluidic flow moves from the direction of right to left).

The enzymatically cleaved DNA can then be collected in the microfluidic output and then quantified via fluorospectrometer by referencing measurements against the calibration curve we derived in ***figure 4.9***. By being careful not to contaminate flow-through from the cell lysis stage of the process, we were able to remain confident that the DNA collected from enzymatic accurately represented only the gDNA tethered upon our pillars. We then repeated the process of cell capture, gDNA extraction, enzyme digestion, off-chip collection, and mass measurement for varying cell counts. Since we wanted to evaluate the sensitivity of our device for the purposes of eventual single cell DNA extractions, we chose to analyze low cell count populations of less than 100 cells. Assuming the human genome contains approximately 3 billion bases, diploid cells would theoretically contain 6.6pg of DNA.[20] This can then be used to evaluate our microfluidic device extraction efficiency by comparing our mass of DNA collected versus original cell count prior to cell lysis. As shown in ***figure 4.11*** below, we obtained a cell mass of 6.7 ±0.2pg per cell across all samples.

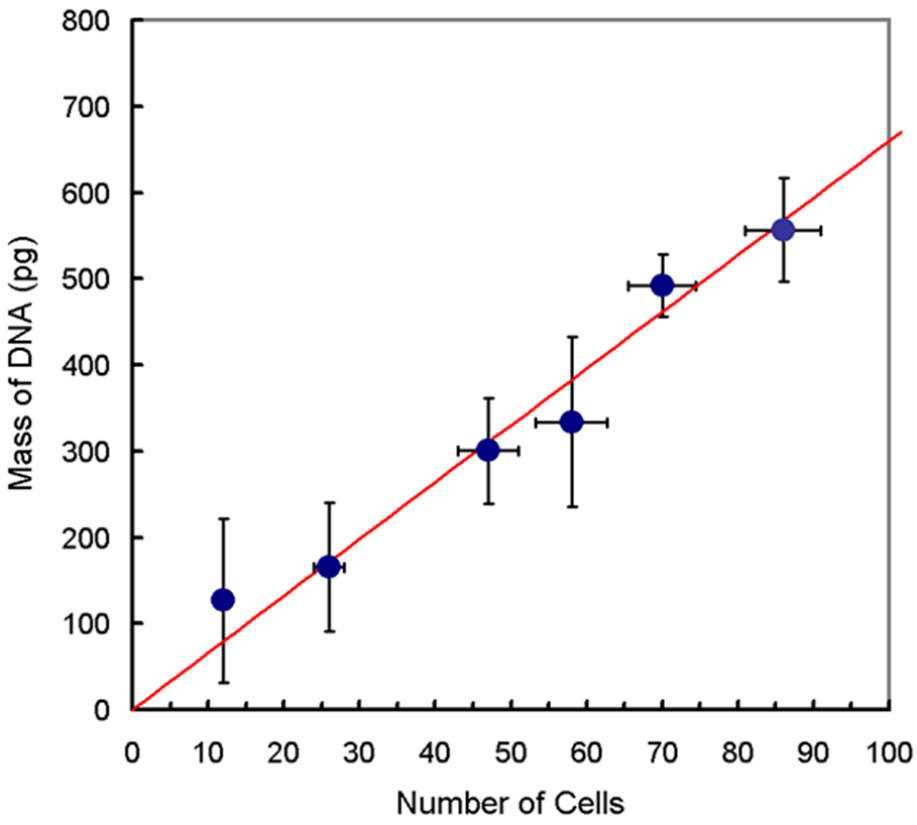

Figure 4.11 – Extracted DNA Mass vs Cell Count

Six different cell counts tested for populations of less than 100 cells. Mass was measured via fluorescence intensity on fluorospectrometer and determined with T4 DNA calibration curve. Horizontal error bars represent dead cells that were present in the micropillar array during cell lysis that could have skewed total DNA mass.

Possible explanations for why our measured values exceed the theoretical value could stem from biological reasons such as that M0-91 may be hyperploid as are many malignant cell types. Another explanation for the mismatch between measured and theoretical DNA mass per cell may be due to errors in cell

counting. In higher cell count samples, many dead cells, identified by their significantly smaller sizes, were often present. Yet, depending on the stage of cell death, DNA may have either completely undergone apoptosis and be too fragmented to become tethered on the micropillars or they may still be relatively in-tact, in which case accounting for their DNA would increase the measured total DNA mass for that given sample. Control experiment using over-cultured dead M0-91 cells showed no DNA capture on micropillar array, confirming DNA fragmentation during apoptosis. Thus, the potential error attributed to dead cell count is represented by the horizontal error bars seen in ***figure 4.11*** above. Here, although the reproducibility of the NanoDrop 3300 measurements were low, leading to a wider range of data for each extraction, our overall assessment is that our microfluidic device is capable of high efficiency DNA capture with little to no DNA loss.

4.3.4 Single Cell DNA Extraction and qPCR

In the realm of single cell analysis, sample loss during DNA extraction or purification equates directly to an incomplete target genome. Although this scenario can be mitigated in population level analysis due to the presence of many genome copies, having at maximum only two copies of each gene in a

single cell requires DNA handling platforms with much higher precision. To evaluate the suitability of our microfluidic device as it applies to single cell analysis, we analyzed single cell DNA capture efficiency with quantitative polymerase chain reaction (qPCR). First, by significantly diluting the concentration of cells prior to loading them into the channel, we were able to load single cells into our microfluidic channel as shown in ***figure 4.12*** below.

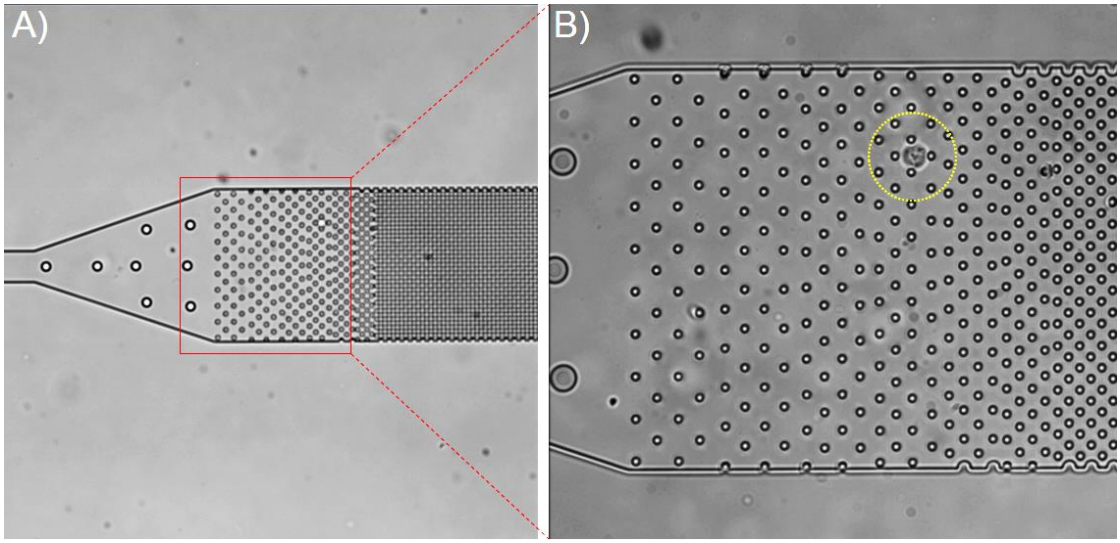

Figure 4.12 – Single Cell Entrapped Within Micropillar Array

(A) Single M0-91 cell lodged within the cell capture region of the micropillar array. (B) Zoom-in of the area cropped in (A) taken with a 20x objective lens highlighting the single cell within the yellow circle. (In all panels, fluidic flow moves from the direction of left to right).

Captured single cells were lysed in our device following identical protocols from the previously described multi-cellular DNA extraction and off-chip collection methods. Bulk-extracted M0-91 gDNA was diluted to three separate DNA mass standards of 3.3pg, 6.6pg, and 9.9pg and then whole genome amplified with qPCR. The resulting amplification curves are shown in ***figure 4.13*** below.

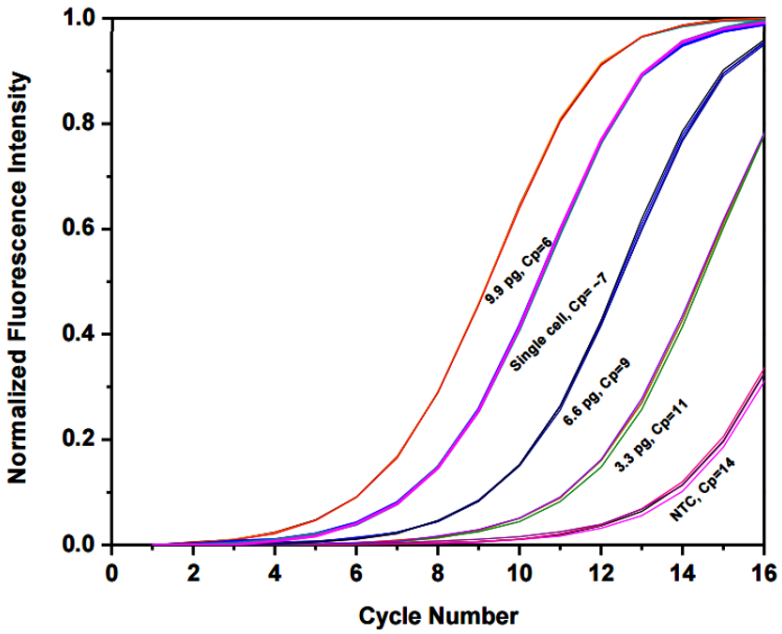

Figure 4.13 – Measured qPCR Fluorescence Intensity versus Cycle Number

C_p value measured for single M0-91 cell DNA on-chip extraction was 7.4. C_p values from DNA standards of 3.3pg, 6.6pg, and 9.9pg were 6, 9, and 11 respectively. Blank control samples with no DNA had a measured C_p value of 14.

Cp values for 3.3pg, 6.6pg, and 9.9pg of DNA was measured to be 6, 9, and 11 respectively. Additionally, qPCR performed on a blank sample of water in the absence of DNA resulted in a C_p of 14. A linear fit was created for C_p value measured versus pre-amplification DNA mass based on the bulk-extracted M0-91 DNA standards and blank sample as shown in ***figure 4.14*** below. The single M0-91 cell on-chip extracted DNA sample C_p measured at 7.4 was then plotted against this linear fit to derive the starting DNA mass prior to amplification.

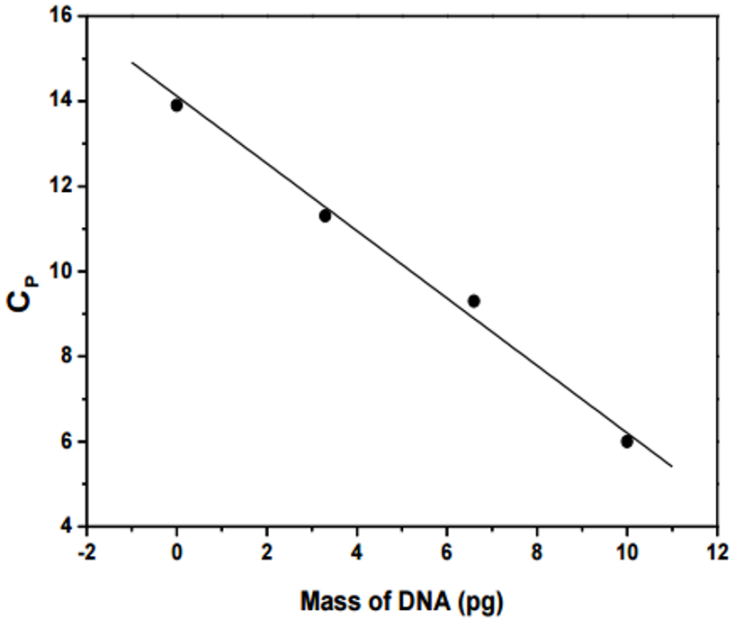

Figure 4.14 – Measured C_p versus Pre-Amplification DNA Mass

Six different cell counts tested for populations of less than 100 cells. Mass was measured via fluorescence intensity on fluorospectrometer and determined with T4 DNA calibration curve. Horizontal error bars represent dead cells that were present in the micropillar array during cell lysis that could have skewed total DNA mass.

In the plot of C_p versus DNA mass, the slope of the linear fit through the DNA standards measured with bulk-extracted DNA was -0.8 ±1 with an R^2 value of 0.988. Through plotting the single cell C_p value of 7.4 along the fit, we determined the original single cell in the microfluidic extraction to have contained 8 ±1pg of DNA. The fact that our measured single cell DNA mass exceeds the theoretical value of 6.6pg could be a result of two factors. The first is that we did not control for cell cycle phase in our experiments and if the single cell that we sampled was in S phase, it is undergoing constant DNA replication in preparation for the G2 phase and cellular division and thus would skew our measured DNA mass.[21] The second factor is that as chromosomal hyperploidy is one of the most common cytogenetic aberrations found in acute myeloid leukemia cells, in which case extra chromosomes present in the original may be the culprit behind the higher than theoretical measured DNA mass.[22-23]

4.3.5 In-Channel H3K9me2 Fluorescence Quantification

Chromatin extraction and imaging in-channel was briefly explored with HeLa cells for the purposes of fluorescence based epigenetic mark quantification. In these experiments, varying cell counts of human cervical cancer HeLa cells were loaded into the microfluidic channel and lysed with a 1% Triton-X buffer. This softer surfactant allowed for permeabilization of the cellular membrane without stripping the nucleosomes from the chromatin as would be the case with using a harsher surfactant such as SDS. As 1% Triton-X leaves the nuclear envelope undissolved, the majority of chromatin within the lysed cells becomes tethered by the nucleus rather than within the micropillar array. This chromatin can then be hybridized with fluorescently tagged antibodies specific to histones and epigenetic histone modifications. In ***figure 4.15*** below, chromatin from HeLa cells lysed with Triton-X are incubated with H3K9me2 specific antibodies before subsequently also being labeled with H3 specific antibodies.

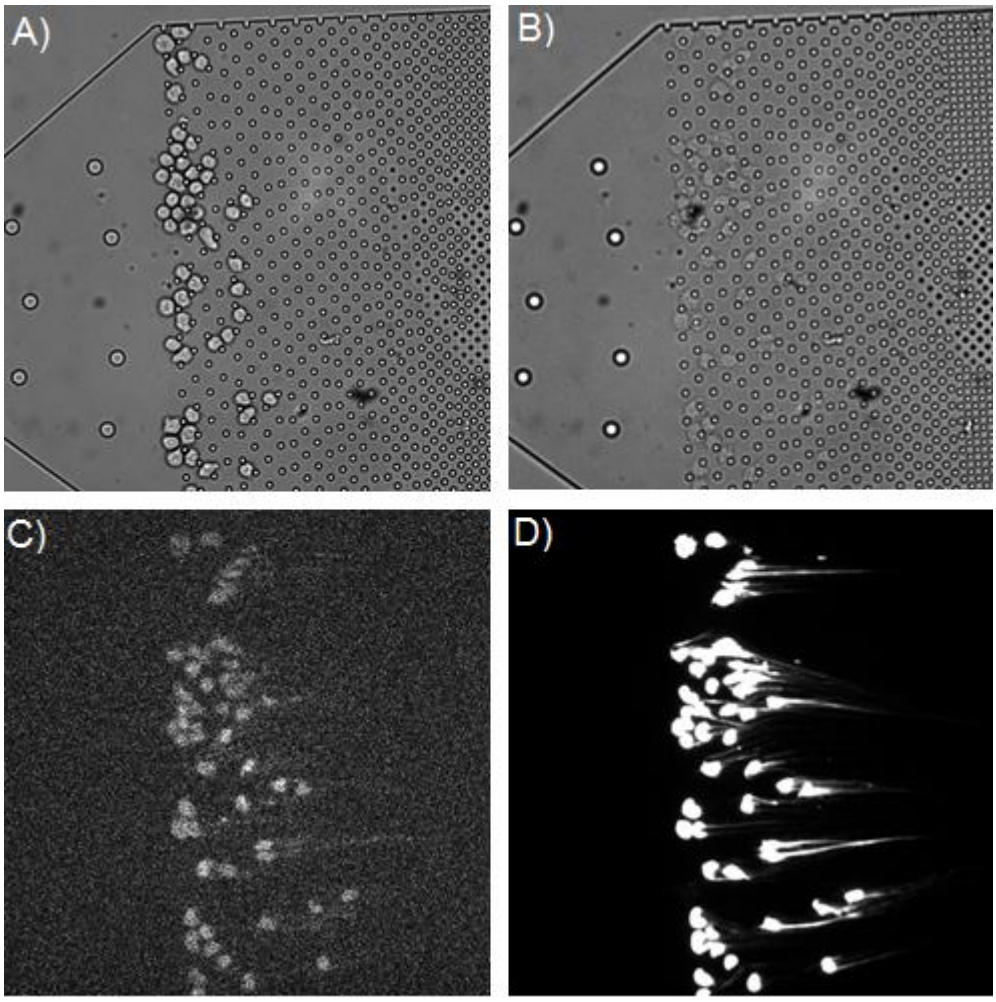

Figure 4.15 – H3 and H3K9me2 Antibody Labeling on Extracted HeLa Chromatin In-Channel

(A) 52 HeLa cells immobilized within the cell capture region of the micropillar array. (B) Post-lysis with 1% Triton-X in 1x PBS. Cell walls have visibility been permeabilized but cellular structures are still visible within the pillar array indicating non-complete dissolving of the nuclear envelope. (C) Chromatin labeled with anti-H3K9me2 antibody tagged with Alexa Fluor 647. (D) After washing out unbound anti-

H3K9me2 antibody, AlexaFluor647-tagged anti-H3 antibody was flowed into the channel to label all histone H3s. (In all panels, fluidic flow moves from the direction of left to right).

The total measured fluorescence intensity values were then corrected for the protein-to-dye ratio which was 1 and 3 for H3K9me2 and H3 respectively. By then taking the corrected fluorescence intensity of H3K9me2 as a fraction of the corrected histone H3 fluorescence intensity, we were able to fluorescently quantify the frequency of H3K9me2 within a group of cells. Furthermore, for cells sufficiently spaced apart, our microfluidic platform allowed us to identify H3K9me2 modification frequency on a single cell basis. Although in our experiments, both histone H3 and H3K9me2 modification were labeled with Alexa Fluor 647, we reason this did not affect fluorescence intensity measurements of either H3K9me2 or histone H3. This is because we labeled the rarer H3K9me2 marks prior to labeling all histone H3 thus incorporating the fluorescence of the H3K9me2 marks into the histone H3 fluorescence measurements. When we took this approach to then measure the H3K9me2 frequency in 75 HeLa cells across 15 experiments, we found an average H3K9me2 frequency of 2.9 ±0.9% which was in accordance with the 2.6 ±0.1% frequency reported by Voigt P. et. al., 2012.[24] These results are shown in ***figure 4.17*** below.

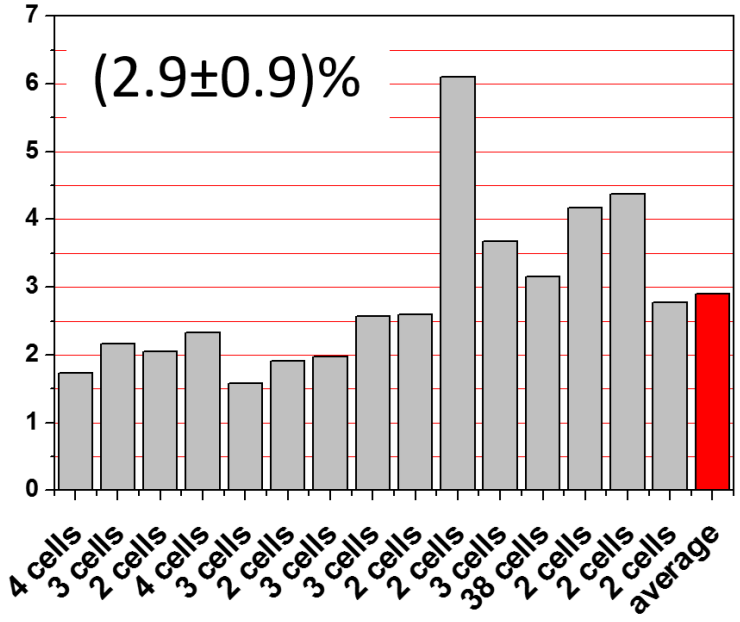

Figure 4.16 – H3K9me2 Frequency versus HeLa Cell Count

Frequency of H3K9me2 as a percentage of total histones per cell measured in 75 cells across 15 separate microfluidic experiments where HeLa wild type cell chromatin was extracted and fluorescently antibody-labeled. Largest cell count performed in a single experiment was 38 HeLa cells, which resulted in a measured H3K9me2 frequency similar to the 75-cell average.

Finally, to confirm that the fluorescence events imaged were a result of target specific antibody binding rather than non-specific dye binding, IgG was conjugated with Alexa Fluor 647 dye and flowed into our device post chromatin

extraction. Results as shown in ***figure 4.18*** validate the specificity of our antibodies and lack of non-specific Alexa Fluor 647 binding events.

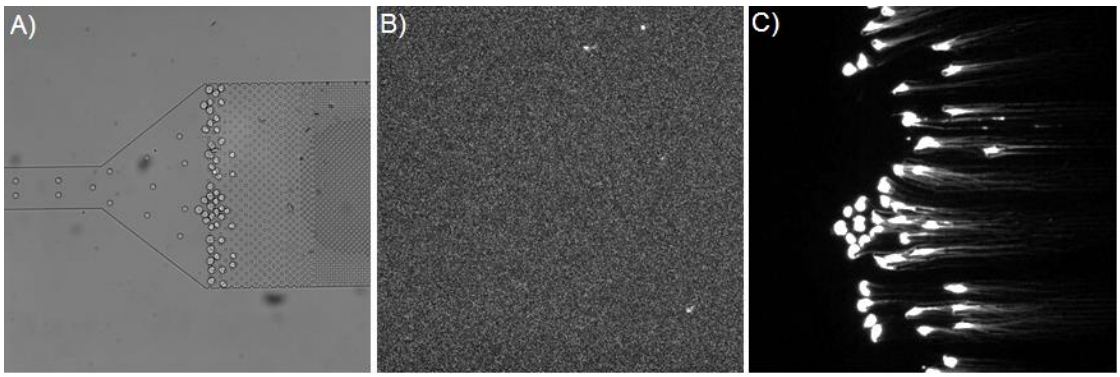

Figure 4.17 – IgG Tagged with Alexa Fluor 647 as a Control for HeLa Chromatin Labeling

(A) 49 HeLa cells immobilized within the cell capture region of the micropillar array. (B) AlexaFluor647-tagged IgG incubated with the extracted chromatin show no visible nonspecific binding. (C) AlexaFluor647-tagged H3 antibody flowed in after washing out IgG shows that all Alexa Fluor 647 staining is specific to the target and not a result of non-specific dye binding. (In all panels, fluidic flow moves from the direction of left to right).

4.4 ALTERNATIVE DEVICE DESIGNS EXPLORED

Several alternative device designs were explored in the process of iterating and improving the performance and function of the device. From these alternative designs we were able to draw conclusions between design parameters and their effects on device performance. On such alternative design explored was a shorter micropillar array region length. In this iteration, the motivation to shortening the pillar array region was for condensing microfluidic channel real-estate. However, as shown in ***figure 4.19***, shorter pillar array regions are unable to contain the entire length of the extracted gDNA from a group of cells. In this case, DNA can be observed to be seen tethered within the tapered region of the channels leading into the output port and in many cases is visible within the output port. DNA was observed to be tethered on the downstream support structures rather than within the micropillar array indicating potential DNA loss during the extraction process as the density of the support structures is much lower than that of the micropillar array.

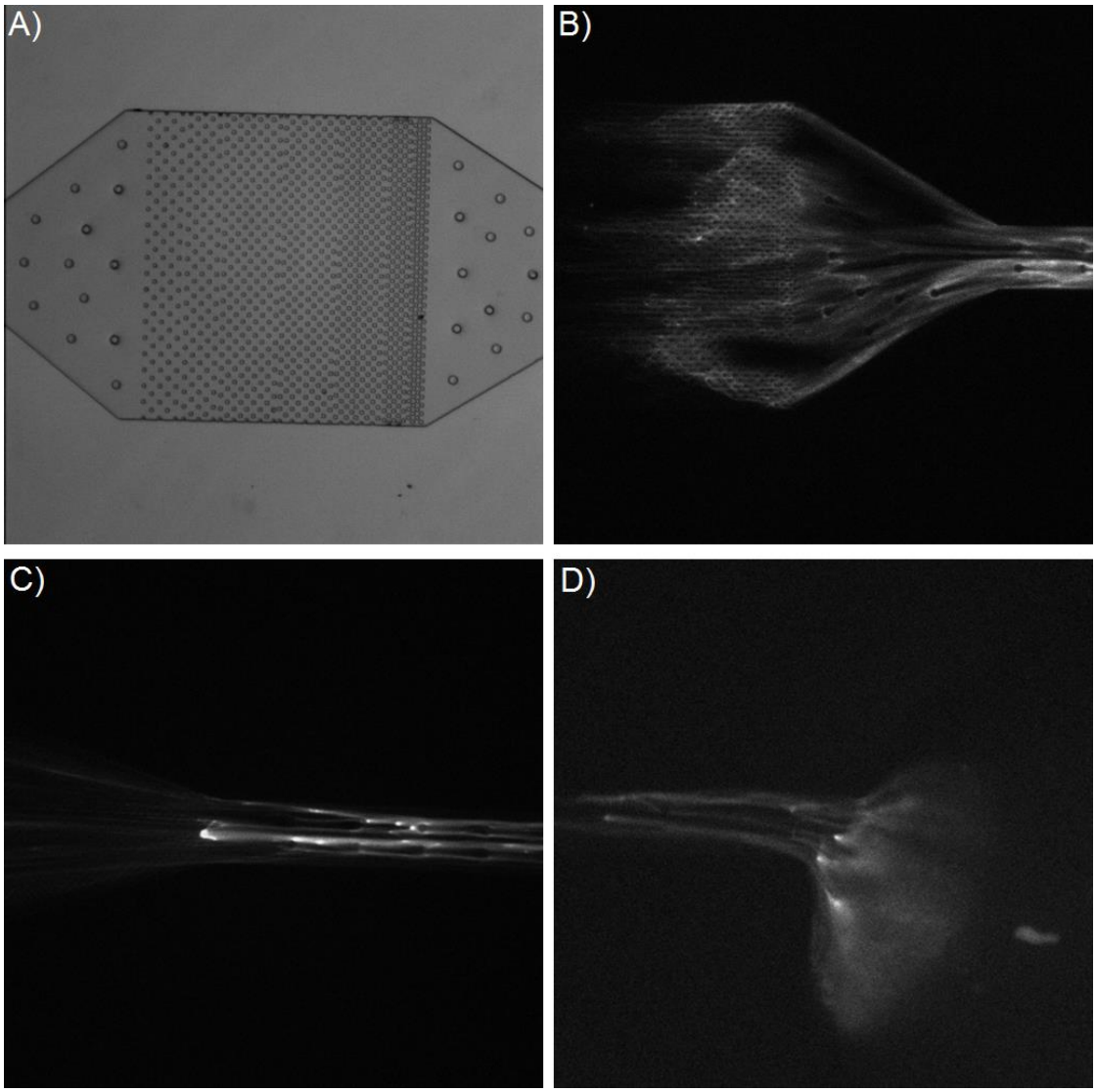

Figure 4.18 – Short Micropillar Arrays Unable to Retain Full Length of Extracted gDNA

(A) Micrograph of a shorter micropillar array region channel geometry. (B) Extracted gDNA from cells loaded into the pillar array region shown in (A). Here, DNA can be clearly seen tethered on support structures beyond the micropillar array. (C) gDNA stained with fluorescent intercalating dye seen extending past the pillar array region

into the tapered region leading to the output port. (D) gDNA seen extending into the output port and upwards in the z-axis into the output tubing. (In all panels, fluidic flow moves from the direction of left to right).

To remediate the short channel lengths, the micropillar array region was extended into a 10cm long region that folded on itself in an 'S' shape. The idea of this design was to allow for the longest human chromosome, chromosome 1 containing 249 million base pairs stretching roughly 8.5cm from end-to-end, to fully fit within the micropillar array unfolded. Unfortunately, this design was found to be unnecessary as the extracted DNA very rarely extended past 1cm. Furthermore, there were challenges in evenly staining the entire length of the pillar array region in the snaking channels. Finally, the large area covered by the micropillar array made small pillar diameters impractical in the CAD design process. In the end, longer micropillar array regions were adopted for the device design, but channels were kept linear in geometry and under 2cm in length.

4.5 CONCLUSION

In this chapter, we described a device for DNA extraction of single or multiple cells by physical entrapment using micropillar arrays. The DNA extraction process uses chemical lysis to dissolve the cellular membrane and nuclear envelope. As the histone proteins are denatured during lysis, the chromosomal DNA can unravel into a lower order structure. In this stage, the genomic DNA traveling downstream with the flow become physically entangled on the downstream micropillars. This entanglement can occur in a number of different configurations, however, as the short length scales of the tethered DNA observed within our micropillar arrays strongly suggest the possibility of multiple tethering points across multiple micropillars for each strand of gDNA as opposed to gDNA tethered on a single micropillar at a single location along each molecule.

The extracted DNA was collected off-chip by in-channel enzymatic release, a process that can be expanded into future integration of our device with downstream DNA processing or analysis platforms. DNA collected from various cell counts were measured via fluorospectrometer. From this, we determined that we are capable of achieving 100% extraction efficiency with the micropillar array. We then used qPCR to analyze our single cell DNA extraction

efficiency and demonstrated that we extracted 8pg of DNA, comparable to the theoretical mass of 6.6pg found in human diploid cells. Finally, we used our device as an imaging platform to quantify the frequency of specific histone modifications by using antibody labeling.

The uniqueness of our single cell DNA extraction device versus existing single cell processing devices lie in our simplicity of design. Our microfluidic device contains no valving, chambering, or peristaltic pumping. DNA tethered within our device can be purified by hydrodynamic flow. We do not rely on chemical surface modification to capture the genomic DNA and we have demonstrated that micropillar arrays can be used to extract DNA from both single cells as well as multi-cells populations.

In the next chapter, we modify our micropillar array device design to increase the throughput 10-fold and tailor the channels for single cell capture. Having shown that we are capable of high efficiency DNA extraction from single cells, we explore isothermal whole genome amplification of tethered genomic DNA extracted from single cells.

4.6 REFERENCES

1. C. Ren, P. He, J. Zhang, Z. Zheng, Y. Qiao, X. Zhao, Cancer Bio. and Therapy, 2011, 11(7), 633-638
2. V. Plaks, C. D. Koopman, Z. Werb, Science, 2013, 341(6151), 1-5
3. J. H. Tsai, J. Yang, Genes and Development, 2013, 27, 2192-2206
4. P. Mehlen, A. Puisieux, Nature Reviews, 2006, 6, 449-458
5. J. Sleeman, P. A. Steeg, European J. of Cancer, 2010, 46(7), 1177-1180
6. M. C. Miller, G. V. Doyle, L. W. M. M. Terstappen, J. of Oncology, 2010, 617421, 1-8
7. B. Kirby, M. Jodari, M. S. Loftus, G. Gakhar, E. D. Pratt, C. Chanel-Vos, J. P. Gleghorn, S. M. Santana, H. Liu, J. P. Smith, V. N. Navarro, S. T. Tagawa, N. H. Bender, D. M. Nanus, P. Giannakakou, PLOS One, 2012, 7(4), e35976, 1-10
8. S. Wang. K. Lui, J. Lui, Z. T. F. Yu, X. Xu, L. Zhao, T. Lee, E. K. Lee, Y. K. Lee, L. W. K. Chung, J. Huang, M. Rettig, D. Seligson, K. N. Duraiswamy, C. K. F. Shen, H. R. Tseng, Angewandte Chemie, 2011, 50, 3084-3088
9. P. Li, Z. S. Stratton, M. Dao, J. Ritz, T. J. Huang, Lab Chip, 2013, 13, 602-609

10. S. Nagrath, L. V. Sequist, S. Maheswaran, D. W. Bell, D. Irimia, L. Ulkus, M. R. Smith, E. L. Kwak, S. Digumarthy, A. Muzikansky, P. Ryan, U. J. Balis, R. G. Tompkins, D. A, Haber, M. Toner, Nature Letters, 2007, 450, 1235-1239
11. S. L. Stott. C. H. Hsu, D. I. Tsukrov, M. Yu, D. T. Miyamoto, B. A. Waltman, S. M. Rothernberg, A. M. Shah, M. E. Smas, G. K. Korir, F. P. Floyd Jr., A. J. Gilman, J. B. Lord, D. Winokur, S. Springer, D. Irimia, S. Nagrath, L. V. Sequist, R. J. Lee, K. J. Isselbacher, S. Maheswaran, D. A. Haber, M. Toner, PNAS, 2010, 107(43), 18392-18397
12. J. Chen, J. Li, Y. Sun, Lab Chip, 2012, 12, 1753-1767
13. R. Boom, C. J. A. Sol, M. M. M. Salimans, C. L. Jansen, P. M. E. Wertheim, J. van der Noordaa, J. of Clinical Microbiol., 1990, 28(3), 495-503
14. A. De, W. Sparreboom, A. van den Berg, E. T. Carlen, Biomicrofluidics, 2014, 8, 054119, 1-11
15. R. Zhang, H. Q. Gong, X. Zheng, C. Lou, C. Sze, Anal. Chem., 2012, 85, 1485-1491
16. T. L. Hawkins, T. O'Connor-Morin, A. Roy, C. Santillian, Nucleic Acids Research, 1994, 22(21), 4543-4544
17. C. W. Price, D. C. Leslie, J. P. Landers, Lab Chip, 2009, 9, 2484-2494
18. J. Wen, C. Guillo, J. P. Ferrance, J. P. Landers, Anal. Chem., 2007, 79, 6135-6142

19. J. Kim, M. Johnson, P. Hill, B. K. Gale, Integr. Biol., 2009, 1, 574-586

20. C. Kohler, R. Radpour, Z. Barekati, R. Asadollahi, J. Bitzer, E. Wight, N. Burki, C. Diesch, W. Holzgreve, X. Y. Zhong, Molecular Cancer, 2009, 8(105), 1-8

21. S. Cooper, Theoretical Biol. And Medical Model., 2006, 3(10), 1-15

22. Luquet, J. L. Lai, C. Barin, L. Baranger, C. Bilhou-Nabera, E. Lippert, C. Gervais, P. Talmant, P. Cornillet-Lefebrve, C. Perot, N. Nadal, M. J. Mozziconacci, M. Lafage-Pochitaloff, V. Eclache, F. Mugneret, C. Lefebvre, C. Herens, F. Speleman, H. Poirel, I. Tigaud, C. Cabrol, P. Rousselot, S. Daliphard, M. Imbert, R. Garand, F. Genevieve, R. Berger, C. Terre, Nature, 2008, 22, 132–137

23. C. Veigaard, J. M. Norgaard, E. Kjeldsen, Cancer Genetics, 2011, 204(9), 516-521

24. P. Voigt, G. LeRoy, W. J. Drury III, B. M. Zee, J. Son, D. Beck, N. L. Young, B. A. Garcia, D. Reinberg, Cell, 2012, 151(1), 181-193

CHAPTER 5

SINGLE CELL WHOLE GENOME AMPLIFICAITON VIA MICROPILLAR ARRAYS

5.1 ABSTRACT

Single cell whole genome amplification (WGA) has long suffered from amplification biases that reduce the accuracy of single cell sequencing data. To this end, an easily adoptable process requiring minimal microfabrication complexity remains desirable. Here, we describe genomic amplification via micropillar array (GAMA) on single human cancer cells from the HeLa cell line. This micropillar array is designed to capture single cells and physically entangle its chromosomal DNA in a fixed position throughout WGA. By testing for the presence of 6 gene loci along the human genome, we demonstrate an improved genome coverage and reduced amplification bias using GAMA as opposed to conventional fluorescence activated cell-sorting (FACS) based single cell assays.

5.2 INTRODUCTION

Single cell analysis has become increasingly important for understanding and diagnosing disease.[1-6] For instance, cellular level aberrations have been shown to play critical roles in tumor heterogeneity, cancer metastasis, drug resistance, and cell fate.[7-12] Investigating these aberrations and differentiating between cell types within a population may give rise to improved treatments, however, single cell handling and analysis remains difficult. Due to having only picogram quantities of DNA, existing workflows cannot sequence single cell genomes directly without amplification due to sensitivity limits.[13-15] Thus, to obtain a sufficient quantity of material for sequencing, single cell WGA is necessary. Among most widely used single cell WGA amplification technique is multiple displacement amplification (MDA), which relies on a combination of random hexamer primers and the strand-displacement properties of the Phi29 polymerase to isothermally amplify DNA.[14,15] However, the primary technical challenge in using MDA for single cell WGA is random amplification bias resulting from chimera formation and non-linear enrichment.[16-18] This bias can be averaged out when analyzing monodisperse multi-cell population samples due to having a multiple copies of each gene. However, biases occurring on the single cell level lead to severe underrepresentation of genome regions that were not amplified early-on in the

MDA reaction.[19] ***Figure 5.1*** below is an example of fractional genome coverage becoming more pronounced as WGA is performed on fewer cells.[31]

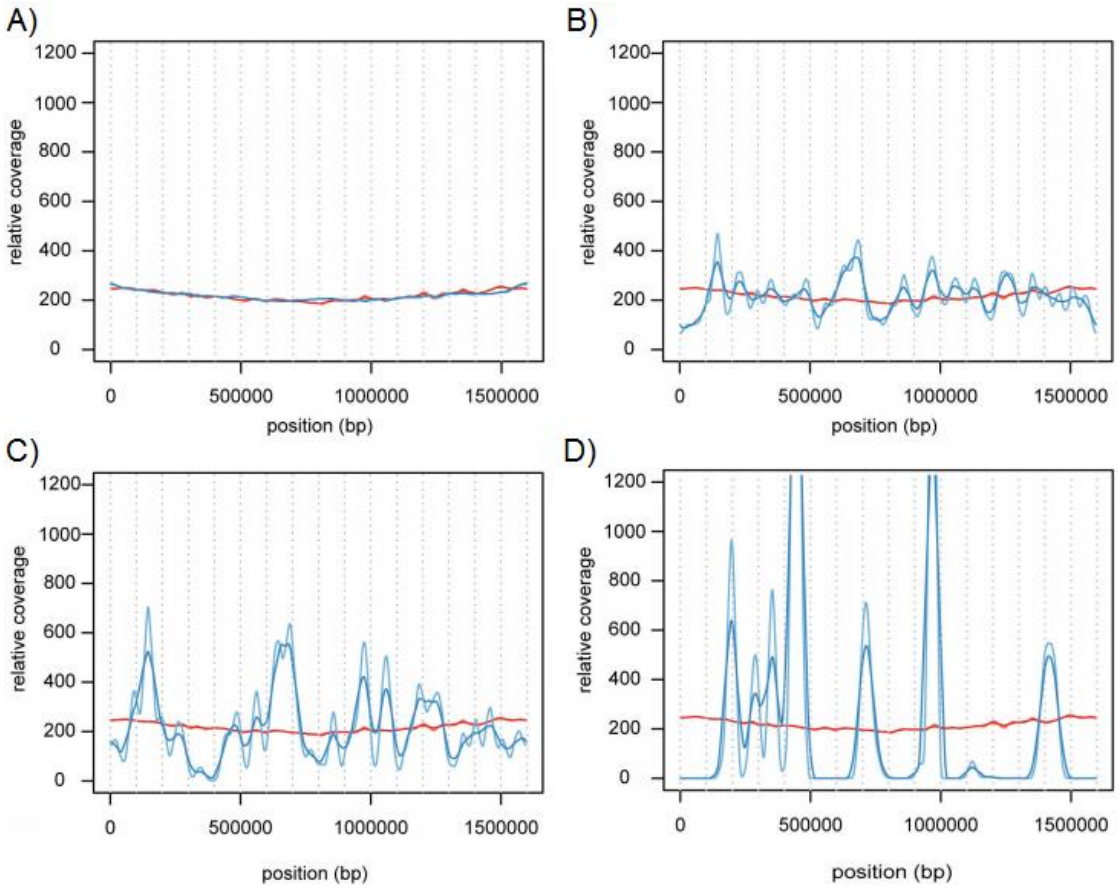

Figure 5.1 – Read Density vs Genome Position from Varying Cell Counts

(Modified from Ellegaard K. M. et. al., 2014)

(A) Read density across position is fairly homogenous when sequencing the DNA amplified from millions of cells. However, this homogenization begins to deteriorate as WGA is performed on fewer and fewer cells with (B) Thousands of cells and (C) 10s of cells. (D) Single cell amplified by WGA shows significant fractional genome coverage.

To this end, several techniques have been found to minimize amplification bias during MDA by reducing reaction volumes.[20] Although the mechanism by which reducing amplification volume reduces bias remains to be fully explained, it has been demonstrated across a number of platforms. These platforms can be broadly categorized into limiting dilution technologies,[21] droplet microfluidic technologies,[22-24] and chambered microfluidic technologies.[25] Limiting dilution technologies provide a high degree of parallelism, but the microwells can suffer from cross-contamination of liquids and reagents.[21] More reliable compartmentalization of single cell genomic material can be achieved via emulsion enclosure and microfluidic chambers, however complex channel geometries and valving systems are required to achieve an integrated platform capable of both single cell isolation and genomic analysis. Hence, exploring alternative methodologies of integrating cell capture and genomic analysis is a critical component of the overall effort to improve single cell sequencing.

Recently, our group has developed a valveless microfluidic device for on-chip single cell capture and DNA extraction.[26] The core of this technology uses micropillar arrays to physically entrap genomic DNA (gDNA) from cells upon

lysis. As this process is purely mechanical, it does not require any chemical modification or cell sample preparation.

Here, we utilized the unique advantages conferred by micropillar arrays as a basis for developing GAMA, a novel microfluidics-based approach towards single cell WGA. GAMA relies on the high capture efficiency and DNA immobilization properties of micropillar arrays to hold template gDNA in a fixed position within the microchannel as reagents for WGA are flowed through. This approach differs fundamentally from previously mentioned technologies in that the template gDNA is subjected to a constant flow throughout the amplification process while the amplified product is washed downstream and collected in the output reservoir. To demonstrate the viability of our approach, we use GAMA to perform MDA-based WGA of single cells and compare the genome coverage, determined by the successful amplification of select gene loci, to conventional assays based on fluorescence activated cell sorting (FACS).

5.3 MATERIALS AND METHODS

5.3.1 Cell Culture

HeLa-GFP cells were cultured in Dulbecco's Modified Eagle medium (DMEM) (Invitrogen) within a T75 flask at 37C and 5% CO2. Cell culture medium was supplemented with 10% fetal bovine serum (FBS) (Atlanta Biologicals; Atlanta, GA), 1% (wt) non-essential amino acids (NEAA) (Gibco, Life Tecnologies), 1% (wt) L-glutamine (Gibco, Life Tecnologies), 2% (wt) HEPES (Quality Biological; Gaithersburg, MD), and 0.1% of 1:100 dilution 2-mercaptoethanol (βME) (Sigma-Aldrich; St. Louis, MO). Cells were passaged at 60% (vl) confluency roughly twice per week.

5.3.2 Device Fabrication

General photolithography fabrication steps were done in accordance with the methods outlined in ***section 5.2.2*** of ***Chapter 5***. However, PDMS was mixed at a ratio of 12:1 base resin to curing agent, rather than 10:1, and was heat cured at a lower temperature 100C for one hour in a Sheldon oven. Also, the design of the single cell microfluidic channels mask was vastly different than that of the original microfluidic pillar array based cell processor. Rather than having a single input port lead to a single output port, we increased the throughput of the device by splitting the output into 10 separate but identical channels.

5.3.3 Single Cell Capture and Lysis

HeLa-GFP cells were trypsinized from T75 flasks with 0.25% Trypsin. Trypsinized cells neutralized with 1:1 dilution of phosphate buffered saline (PBS) buffer, spun down in a centrifuge, and then resuspended in fresh PBS at a concentration of 1:50. The cell suspension was flowed into the microfluidic device via pressure driven flow at 2 psi with bone-dry nitrogen gas (Airgas; Radnor Township, PA). The infusion apparatus was then disconnected from the microfluidic device's input port, washed with alternating cycles of 100% ultrapure water (Invitrogen; Carlsbad, CA) and 100% ethanol to remove the remaining cells within the reservoir, and then reconnected to the microfluidic device input port. Sterile PBS buffer was then flowed into the microfluidic device for 5 minutes to allow uncaptured cells to either be captured within the cell capture region or to flow through the device into the output reservoirs. The output reservoirs of the device were then emptied and rinsed with 100% ultrapure water.

Lysis buffer comprised of 6M guanidinium thiocyanate (Sigma-Aldrich; St. Louis, MO) in water was flowed into the microfluidic device for 5 minutes also

by pressure driven flow at 2 psi. After visually confirming cell lysis in all ten channels, the lysis buffer was removed from the input reservoir and the reservoir rinsed with 100% ethanol before flushing the entire microfluidic device with 100% ethanol for 5 minutes. The ethanol is replaced by washing with 100% ultrapure water for 5minutes and then finally replaced by PBS buffer. The output reservoirs that now contain a mixture of cell lysates, lysis buffer, ethanol, water, and PBS was then emptied and cleaned via rinsing first with 100% ethanol and then 100% ultrapure water. The genomic DNA tethered within the microfluidic device is now ready for whole genome amplification.

5.3.4 On-Chip Whole Genome Amplification

Whole genome amplification (WGA) of the single cell genomic DNA tethered within the micropillar array region of the microfluidic device was carried out using reagents from the REPLI-g UltraFast Mini Kit (Qiagen; Hilden, Germany). Prior to starting the reaction, 280ul of buffer D1 was made by adding 35ul of buffer DLB to 245ul of ultrapure H2O. 400ul of buffer N1 was then prepared by adding 40ul of stop solution to 360ul of ultrapure H2O. Finally, 288ul of master mix was made by adding 18ul of polymerase to 270ul Repli-G UltraFast reaction buffer.

To denature the double stranded gDNA tethered on the micropillar array, buffer D1 was flowed through the device continuously at room temperature for 8 minutes. Buffer D1 was then removed and the device was flushed with buffer N1 for 15 minutes. Afterwards, both the infusion apparatus and the ten output reservoirs were emptied and washed with 100% ethanol and then 100% ultrapure water. The infusion apparatus was then loaded with the master mix solution and pressure was dialed down to 0.5 psi. Pressure was then held constant throughout the entire duration of the 3.5 hour reaction amplification reaction while the device was placed atop a hot-plate set to 33C. After the reaction was completed, 5ul of ultrapure H2O was added to each output reservoir. Each output reservoir was then pipette mixed and the solution containing amplified genomic DNA was collected off-chip into a polymerase chain reaction (PCR) tube. Each output reservoir was then rinsed once more with 10ul of ultrapure H2O and the rinse was collected into the respective PCR tube containing the amplified product. All samples were placed in a -20C freezer until further use.

5.3.5 FACS Single Cell WGA

A FACS machine (Becton Dickinson Biosciences; San Jose, CA) was used to sort single HeLa-GFP cells into a PCR-compatible microwell plate (Bio-Rad; Hercules, CA) with each well containing 5ul of sterile PBS buffer. The microwell plate was then spun down in centrifuge at 1000g for 5 minutes to ensure that sorted single cells were sitting at the bottom of their respective wells. Buffer D2 and master mix were then prepared according to the Repli-g UltraFast kit's protocol. To lyse the single cells in each microwell, 5ul of buffer D2 was added to each well and incubated on ice for 10 minutes. 5ul of stop solution was then added to each well and incubated on ice for 5 minutes. Finally, 53.3ul of master mix was added to each well and the microwell plate was placed in a thermocycler (Eppendorf; Germany) set to hold at 30C for 3.5 hours.

5.3.6 Gene Loci PCR

Primers were designed to target 150bp-200bp regions within six gene loci (ERBB2 17q12, PRMT2 21q22, P53 17p13, CCND1 11q13, TRAM1 8q13, and MyC8q24) and ordered through Integrated DNA Technologies (IDT; Coralville, IA). Lyophilized primers were dissolved in water to a concentration of 10μM.

Then, following the protocols from the Taq DNA Polymerase Kit (Life Technologies; Carlsbad, CA), 50ul reaction were prepared for each of the 6 gene loci for every collected single cell WGA sample. 30 cycles of PCR were carried out and the PCR product was run on a 2.3% agarose gel via electrophoresis. Using a 2-log ladder (New England Biosciences; Ipswich, MA), the appropriate size region of 100bp-200bp was evaluated for the presence or absence of the gene.

5.4 RESULTS

5.4.1 Channel Design and Experimental Setup

Figure 5.2 shows the overall experimental setup for GAMA. To create the chip device, a slab of mold-casted PDMS (polydimethylsiloxane) imprinted with the channel geometry is bonded to a glass slide to create the microfluidic device. Reagents are loaded into the device via pressure driven flow from an infusion apparatus housing a large fluid reservoir. Fluid that is loaded into the infusion reservoir can be easily exchanged and replaced via pipetting. The infusion apparatus is a two-part mechanism consisting of a reservoir portion that can be

connected to the PDMS and a cap that is connected to the nitrogen gas cylinder used to drive channel flow.

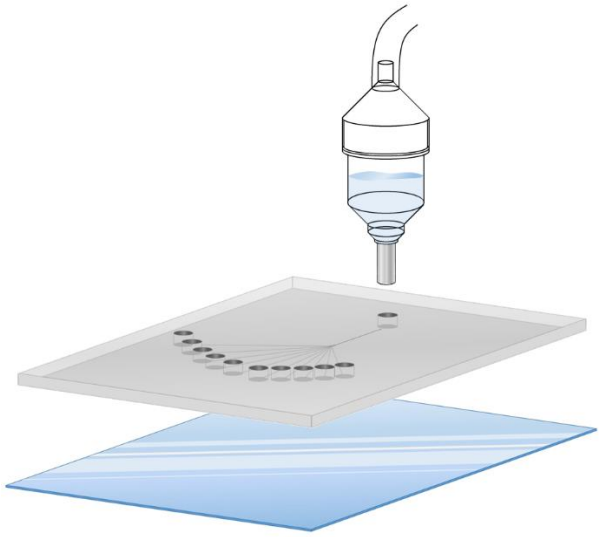

Figure 5.2 – GAMA Experimental Setup Overview

Mold-casted PDMS containing a single input port and 10 parallel output ports are bonded to glass silica via oxygen-plasma bonding. An infusion apparatus is connected to the input port to provide pressure driven flow of desired fluids. Liquid can be manually loaded into the infusion apparatus at will.

Figure 5.3 shows a top down view of the device design. The GAMA device has a single input port and 10 separate output ports allowing multiple single cell samples to be run in parallel. These ten channels each contain identical designs

consisting of a single cell capture region and micropillar array ***figure 5.3(A)***. To show the device in scale, ***figure 5.3(B)*** shows four such devices can be casted from a 4-inch silicon wafer mold as a single slab and bonded to a glass-silica wafer. Experiments were run with the microfluidic device mounted on the stage of an Olympus IX-70 inverted microscope (Olympus; Center Valley, PA) to image and observe the microfluidic channels in real time.

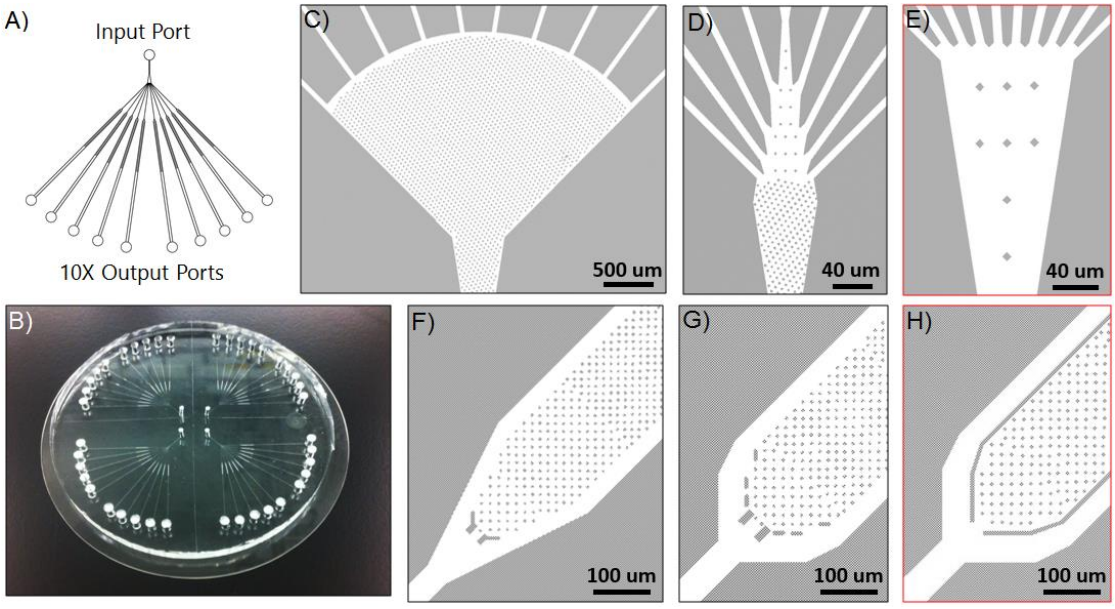

Figure 5.3 – 10-Channel Device Geometry and Design

(A) Device schematic showing valveless 10-channel device design. (B) Picture taken of 4 separate 10-channel devices made from a single PDMS slab bonded to a 4-inch diameter glass silica wafer. (C-E) Chip design variations explored for the bifurcation

point with (E) being the final design used. (F-H) Chip design variations explored for the individual channel cell capture region explored with (H) being the final design used.

5.4.2 Single Cell Whole Genome Amplification On-Chip

As depicted by the graphic in ***figure 5.5(A)***, the single cell capture region consists of a series of posts arranged in an orientation to allow only a single cell to be arrested in the apex of the micropillar array. Barriers surrounding the micropillar array prevent non-arrested cells from lodging themselves in the micropillar array. Upon introduction of lysis buffer, the micropillar array will physically immobilize the gDNA. This immobilization process occurs as a result of the chromosomal DNA being physically entangled on the pillars due to their centimeter scale lengths, while smaller cellular debris such as lipids, proteins, RNA, and mDNA (mitochondrial DNA) are washed away downstream. The immobilized gDNA can be imaged via fluorescent staining with DNA intercalating dye labels such as with YOYO-1 in ***figure 5.5(B)***.

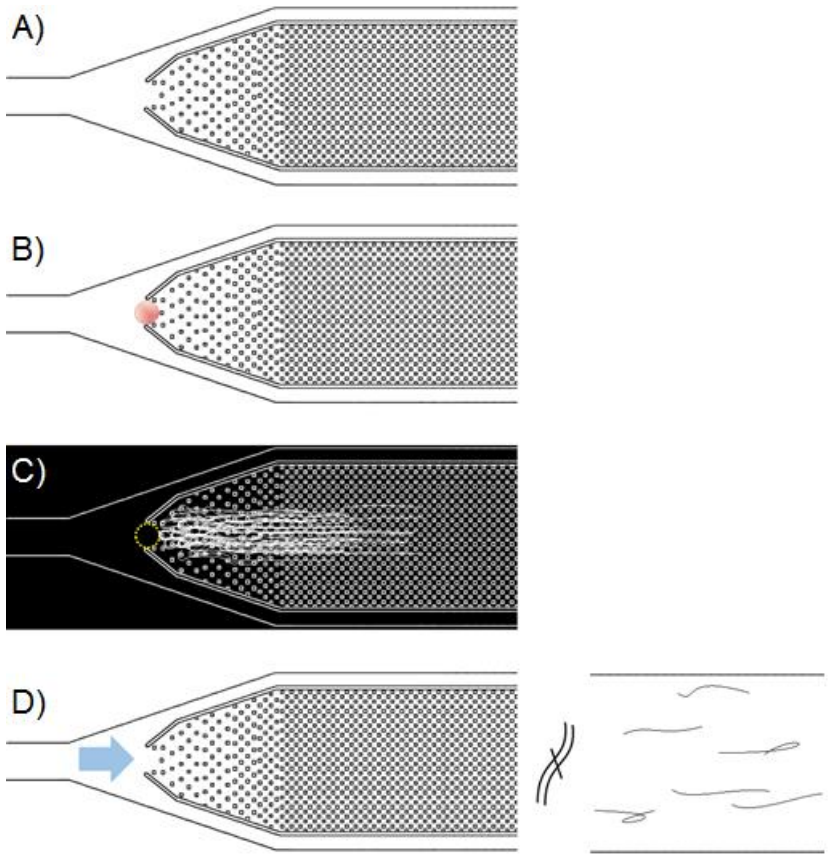

Figure 5.4 – GAMA Process Workflow Illustration

(A) Flow moves from left to right in this schematic and fluids can be exchanged within the input reservoir to control the local environment within the channels. (B) Artist depiction of a single cell that is trapped within the apex of the micropillar array. Side walls enclosing the micropillar array prevent additional cells flowing into the channel from having their DNA immobilized within the pillar array. (C) Upon cell lysis, genomic DNA from the trapped cell will become entangled in downstream pillars and can be visualized via fluorescence staining. (D) Reagents for whole genome amplification is flowed into the channel. As amplification occurs, product strands

elongate originating from the template, but as they depart from the template genomic DNA, they are washed downstream and collected in output reservoirs.

To perform single cell GAMA on the immobilized gDNA tethered within the microfluidic chip, we used MDA with reagents from the commercially available Repli-G UltraFast Mini Kit (Qiagen). An illustration of the GAMA workflow can be seen in ***figure 5.4*** above. After cell capture and lysis, ***figure 5.4(B)*** denaturation buffer D1 was prepared according to Repli-G kit protocols and flowed into the microfluidic device to cleave the hydrogen bonds of the double stranded gDNA into single stranded DNA. Buffer D1 was then neutralized by flushing the channel with neutralization buffer N1. Finally, the master mix containing bases and polymerase was introduced to the channels and the microfluidic chip was set atop a heated hot plate to initiate the amplification reaction. Although we had initial concerns that denatured DNA molecules would rapidly reanneal during the amplification step due to being suspended in close proximity to complimentary strands, we saw no evidence of this reannealing behavior occurring.

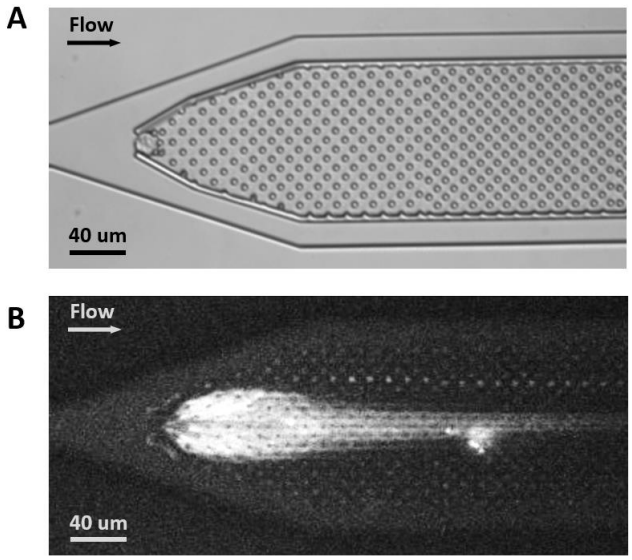

Figure 5.5 – Single Cell Capture and DNA Extraction

Micrograph showing single cell capture (A), and subsequent lysed cell imaged under fluorescence with YOYO-1 intercalating dye staining of genomic DNA immobilized within the pillar array region (B).

Another concern was that the highly branched structures characteristic of isothermal amplification with Phi29 would occlude our channels, however we did not observe any buildup or clogging of the device. We reason that this is because the average sized fragments produced from the MDA, roughly 12kb in length, are too small to wrap around the 1.5μm diameter PDMS micropillars without slipping off. Furthermore, the amplification reaction occurs under a constant flow, thus, as soon as the amplified fragment is detached from its

template strand, it is carried downstream along established flow lines into the corresponding output reservoir.

The amplified DNA can then be collected from the output reservoirs and the gDNA tethered within the micropillar array can be washed and amplified again in further rounds. The advantage conferred from multiple rounds of amplification using GAMA is that regions of the genome that are randomly overrepresented in one amplification will not carry into the next amplification. This is because random amplification bias occurs as a result of the exponential growth in the number of fragments in the reaction as the reaction progresses. Thus, regions of the genome that are amplified first will quickly pull ahead in representation over regions of the genome that are not amplified until later in the reaction. With GAMA, the amplified product can be collected while the template gDNA is still retained within the channels thereby "resetting" the molecule count of the next amplification round. Averaged over multiple amplification rounds, GAMA would theoretically remove random amplification bias and improve genome coverage. As random amplification bias is inherent in single cell MDA, GAMA is a necessary process to obtain full genome coverage.

5.4.3 Validation and Gene Loci Detection

For the reason that MDA has been shown to produce non-specific product in a prolonged reaction, simply quantifying the amount DNA collected from the output reservoirs is insufficient to determine the success or failure of the on-chip single cell GAMA process. To differentiate DNA amplified from gDNA versus non-specific product, 6 different gene loci in the human genome were selected to act as sampling intervals. Using the product collected from GAMA as a template for PCR, the presence or absence of each of the 6 gene loci was evaluated as a means of assessing the bias and overall genome coverage of GAMA.

Figure 5.7 shows the number of gene loci detected from six single cells amplified with GAMA as well as the number of gene loci detected from bulk (107 cells) using the same Repli-G UltraFast kit. On-chip negative controls with 0 cells expectedly did not amplify any gene loci. Furthermore, as the two main sources of potential contamination that may occur are (1) unaccounted cells trapped being within the microfluidic channel and (2) off-chip sample handling, we ran an experiment where a single device had channels containing both a single cell as well as a negative no-cell control. ***Figure 5.6*** shows micrographs taken from the cell capture region of the 10 channels in a single device. Channels containing

a single HeLa-GFP cell (2, 5, and 7) were compared to channels with 0 cells (1, 3, and 10) in the number of gene loci detected post GAMA. While it was found that many of the gene loci were present in the GAMA product collected from single cell channels, 0 gene loci were detected in empty channels on the same device. This result eliminates the possibility of on-chip contamination being a contributing factor in gene loci detection. Samples from channels such as channel 4, 6, 8, and 9 are disregarded due to having multiple cells. Factors contributing to the capture of multiple cells within a channel are that the channel dimensions and micropillar spacing need to be further optimized for the specific cell type being used and rarely, cells become adhered onto the glass surface within the device due to non-specific binding. In future iterations of the device, non-specific cell adherence can be prevented though treating the glass surface with blocking agents or charge-shielding the channel.

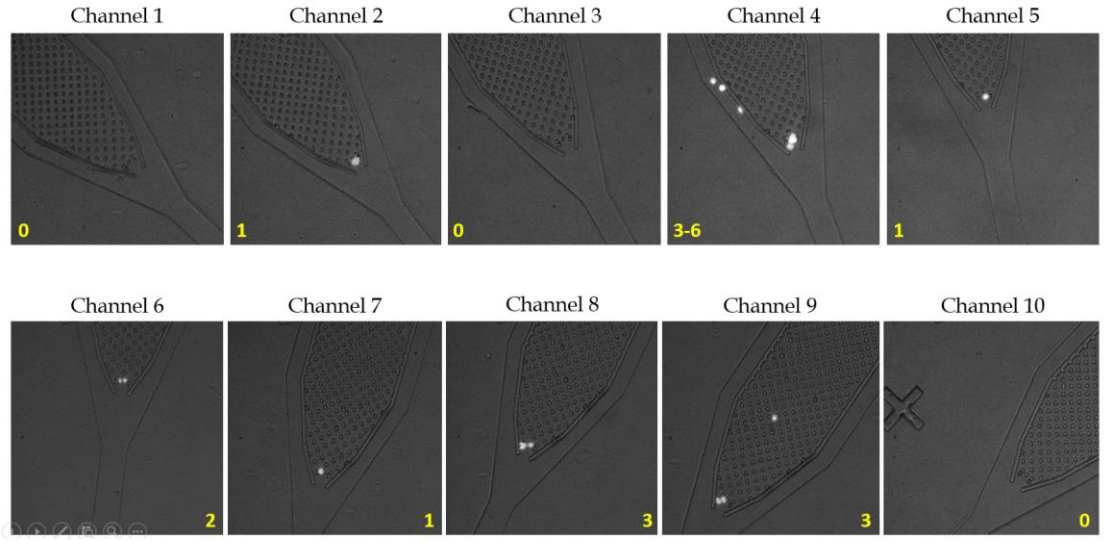

Figure 5.6 – HeLa-GFP Cell Capture in 10-Channel Device

Compiled series of micrographs taken from the cell capture region of a single 10-channel device. Channels containing single cells (2, 5, &7) are analyzed in comparison to empty channels (1, 3, &10), which serve as negative controls from the same device.

5.4.4 Single Cell WGA with FACS

To compare our on-chip single cell GAMA results to single cell WGA in absence of a micropillar array, we performed Fluorescence Activated Cell Sorting (FACS) to isolate various numbers of cells into PCR-compatible 96-well plates and amplified the gDNA from these cells using the same reaction times and reagents. Our findings, shown in ***figure 5.7***, is that when maintaining the same

WGA parameters as GAMA, FACS isolated single cells only amplified one or two gene loci compared to the 4 to 6 amplified by the GAMA process. Finally, past work has observed MDA to exhibit random bias behaviors on the single cell level but non-random bias on a multi-cell level.[13,27-30] Our FACS results support this claim as we have observed that biases occurring in samples of 25 cells consistently underrepresent the same gene loci whereas single cell bias showed no such pattern.

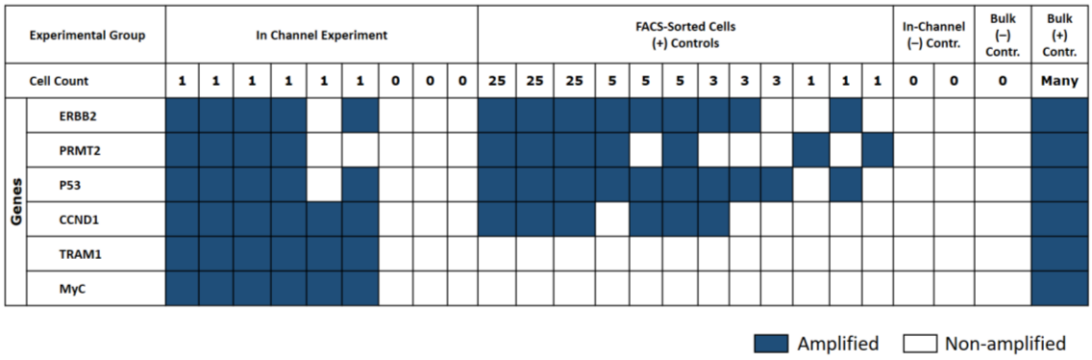

Figure 5.7 – Gene Loci Detection of In-Channel versus FACS Single Cell WGA Compared with Positive and Negative Controls

Compiled table of genome coverage analyzed by detection of 6 cancer-relevant gene loci. Using GAMA, single cell WGA reflects up to 6/6 gene loci detected versus up to 2/6 gene loci detected using conventional single cell WGA via FACS. In-channel single cell

negative controls with 0 cells expectedly show no gene loci coverage and off-chip bulk-level analysis affirms the specificity of primers used in loci detection.

5.5 CONCLUSION

We have described a micropillar-based microfluidic device capable of on-chip single cell processing and WGA. Unlike conventional single cell platforms, GAMA is capable of physically separating the template gDNA from the amplified product during WGA as well as controlling the fluid environment surrounding the gDNA. This property allows GAMA to be used in overcoming random amplification bias such as when performing single cell WGA with MDA. Here, we have taken the first step to demonstrate a reduced amplification bias for single cell WGA using GAMA. This was accomplished by showing that we could reliably amplify more gene loci of the genome from single HeLa cells using GAMA as opposed to single cells isolated through FACS. In future work, we envision being able to use GAMA for multiple rounds of amplification performed in series on a single genome template. Doing so would reset the product pool molecule count at each intermediate washing step, thereby resetting the amplification bias for each round. Subsequently, compiling the randomly over-represented regions of each amplification round may serve as a

means to improve total genome coverage in bias-vulnerable amplification scenarios such as single cell WGA.

5.6 REFERENCES

1. K. R. Chi, Nature Methods, 2014, 11, 13-17
2. E. Shapiro, T. Biezuner, S. Linnarsson, Nature Reviews Genetics, 2013, 14, 618-630
3. X. Zhang, S. L. Marjani, Z. Hu, S. M. Weissman, X. Pan, S. Wu, Cancer Research, 2016, 76(6), 1305-1312
4. N. E. Navin, J. Hicks, Genome Medicine, 2011, 3(31), 1-12
5. K. Galler, K. Brautigam, C. Grobe, J. Popp, U. Neugebauer, Analyst, 2014, 139, 1237-1273

6. C. E. Sims, N. L. Allbritton, Lab Chip, 2007, 7, 423-440

7. X. Liu, F. Long, H. Peng, S. J. Aerni, M. Jiang, A. Sanchez-Blanco, J. I. Murray, B. Mericle, S. Batzoglou, E. W. Myers, S. K. Kim, Cell, 2009, 139(3), 623-633

8. G. Guo, M. Huss, G. Q. Tong, C. Wang, L. L. Sun, N. D. Clarke, P. Robson, Developmental Cell, 2010, 18(4), 675-685

9. P. Dalerba, T. Kalisky, D. Sahoo, P. S. Rajendran, M. E. Rothenberg, A. A. Leyrat, S. Sim, J. Okamoto, D. M. Johnston, D. Qian, M. Zabala, J. Bueno, N. F. Neff, J. Wang, A. A. Shelton, B. Visser, S. Hisamori, Y. Shimono, M. Wetering, H. Clevers, M. F. Clarke, S. R. Quake, 2011, 29(12), 1120-1130

10. Powell, A. H. Talasaz, H. Zhang, M. A. Coram, A. Reddy, G. Deng, M. L. Telli, R. H. Advani, R. W. Carlson, J. A. Mollick, S. Sheth, A. W. Kurian, J. M. Ford, F. E. Stockdale, S. R. Quake, R. F. Pease, M. N. Mindrinos, G. Bhanot, S. H. Dairkee, R. W. Davis, S. S. Jeffrey, PloS One, 2012, 7(5), e33788

11. G. Deng, S. Krishnakumar, A. A. Powell, H. Zhang, M. N. Mindrinos, M. L. Telli, R. W. Davis, S. S. Jeffrey, BMC Cancer, 2014, 14:456

12. L. Paguirigan, J. Smith, S. Meshinchi, M. Carroll, C. Maley, J. P. Radich, Science Translational Medicine, 2015, 7(281), pp. 281re2

13. C. F. de Bourcy, I. D. Vlaminck, J. N. Kanbar, J. Wang, C. Gawad, S. R. Quake, PloS One, 2014, 9(8), e105585

14. N. E. Navin, Genome Biology, 2014, 15:452

15. R. S. Lasken, Biochem Soc Trans, 2009, 37(2), 450-453

16. R. S. Lasken, T. B. Stockwell, BMC Biotech, 2007, 7:19

17. S. Rodrigue, R. R. Malmstrom, A. M. Berlin, B. W. Birren, M. R. Henn, S. W. Chisholm, PloS One, 2009, 4(9), e6864

18. M. Chen, P. Song, D. Zou, X. Hu, S. Zhao, S. Gao, F. Ling, Plos One, 2014, 9(12), e114520

19. Y. Yang, H. S. Rho, M. Stevens, A. G. Tibbe, H. Gardeniers, L. W. Terstappen, Lab Chip, 2015, 15, 4331-4337

20. C. A. Hutchison, H. O. Smith, C. Pfannkoch, J. C. Venter, PNAS, 2005, 102(48), 17332-17336

21. J. Gole, A. Gore, A. Richards, Y. J. Chiu, H. L. Fung, D. Bushman, H. I. Chiang, J. Chun, Y. H. Lo, K. Zhang, Nature Biotech, 2013, 31(12), 1126-1134

22. M. Sidore, F. Lan, S. W. Lim, A. R. Abate, Nucleic Acids Research, 2015, 1-9

23. Y. Fu, C. Li, S. Lu, W. Zhou, F. Tang, X. S. Xie, Y. Huang, PNAS, 2015, 12(38), 11923-11928

24. Y. Nishikawa, M. Hosokawa, T. Maruyama, K. Yamagishi, T. Mori, H. Takeyama, Plos One, 2015, 10(9), e0138733

25. Y. Marcy, T. Ishoey, R. S. Lasken, T. B. Stockwell, B. P. Walenz, A. L. Halpern, K. Y. Beeson, S. M. Goldberg, S. R. Quake, Plos Genetics, 2007, 3(9), 1702-1708

26. J. J. Benitez, J. Topolancik, H. C. Tian, C. B. Wallin, D. R. Latulippe, K. Szeto, P. J. Murphy, B. R. Cipriany, S. L. Levy, P. D. Soloway, H. G. Craighead, Lab Chip, 2012, 12, 4848-4854

27. C. Z. Zhang, V. A. Adalsteinsson, J. Francis, H. Cornils, J. Jung, C. Maire, K. L. Ligon, M. Meyerson, J. C. Love, Nature Communications, 2015, 6 (6822), 1-10

28. Y. Fu, C. Li, S. Lu, W. Zhou, F. Tang, X. S. Xie, Y. Huang, PNAS, 2015, 112(38), 11923-11928

29. Raghunathan, H. R. Ferguson, C. J. Bornarth, W. Song, M. Driscoll, R. S. Lasken, App Environmental Microbiology, 2005, 71(6), 3342-3347

30. K. Zhang, A. C. Martiny, N. B. Reppas, K. W. Barry, J. Malek, S. W. Chrisholm, G. M. Church, Nature Biotech, 2006, 24(6), 680-686

31. K. M. Ellegaard, L. Klasson, S. G. Andersson, PloS One, 2013, 8(11), e82319

CHAPTER 6

CONSLUSION AND FUTURE WORKS

6.1 CONCLUSION

Technologies in genetic and epigenetic analysis of DNA and chromatin have moved away from relying on indirect methods of analyzing amplified genetic material into the realm of directly interrogating the DNA or chromatin molecule of interest itself. Most commonly achieved with imaging, platform technologies built for the purposes of fluorescence based mapping are often low throughput or incompatible with chromatin. Here, we have developed a sequence independent high-throughput technique of molecular combing as applied to native chromatin and demonstrated labeling and imaging of epigenetic histone modifications. However, this technology relies on having many nanograms of material for the combing process, making it impractical for cases of rare of single cell analysis. To this end, we have designed and developed a micropillar array based microfluidic device and a novel WGA approach, GAMA.

This microfluidic device and the related GAMA process is unique in its ability to retain the original genomic DNA from single cells post cell lysis. Unlike existing single cell technologies the GAMA process is fundamentally different by three key factors. The first key factor is that unlike microwell based technologies, microfluidic chamber technologies, and droplet microfluidic technologies where cell lysis and DNA amplification occurs within a single compartment, cell lysate is flushed from the microfluidic channels prior to DNA amplification. This amounts to a lower concentration of cellular debris encountered by amplification reagents during WGA. Furthermore, sequence-dependent amplification bias resulting from exponential amplification can be essentially reset by washing the channel of all amplified material after a round of amplification and then initiating a new round of amplification. By doing so, certain sequences that are exponentially more represented are unable to affect future rounds of amplification. As single cell WGA bias has been observed to be random, multiple rounds of WGA on the tethered gDNA from a single cell can eventually have the various sequences from each round pooled to overcome underrepresented regions of the genome.

The second differentiating factor between GAMA and existing single cell technologies is that GAMA performs WGA under constant hydrodynamic flow, where the product is washed downstream and separated from the location of

the gDNA that is tethered within the pillar array. Here, fluid flow containing polymerase, bases, and primers necessary for DNA amplification are constantly replenishing the local environment surrounding the tethered DNA in a mode of amplification previously unstudied.

Finally, the third differentiating factor stems from the configuration of DNA in GAMA versus other techniques. In GAMA, the constant hydrodynamic flow during amplification drives tethered DNA molecules within the micropillar array to adopt a linearized configuration as opposed to a Brownian-motion derived configuration. As such, the gDNA tethered within micropillar arrays are more uniformly accessible to polymerases in solution compared to coiled DNA in free solution. This increased accessibility may be the reason we observed improved genome coverage of single cell WGA using GAMA versus single cells that were isolated by FACS and amplified in a 96-well plate.

6.2 WHOLE GENOME MULTIPLE AMPLIFICATION

The power of being able to perform multiple rounds of WGA using GAMA affords us the ability to study many previously unanswerable questions. The first of these questions is whether or not single cell WGA bias is truly random.

Currently, amplification bias in single cell whole genome sequencing (WGS) occurs as a result of allele dropout during WGA, false sequencing, chimera formation, and exponential amplification.[1-6] Although false positives in base calling during sequencing can be fixed with higher sequencing depth, allele dropout, chimeras, and sequence-dependent amplification bias is an inherent part of the current single cell WGA methodologies.

In 2014, deBourcy et. al. found that MDA amplification bias is a function of reaction gain, where amplifications beyond 10^6 results in lower genome coverage as shown in ***figure 6.1***.[5] In a separate experiment, the authors performed a second round of WGA on the product of the first round, and find even more pronounce fractional genome coverage.

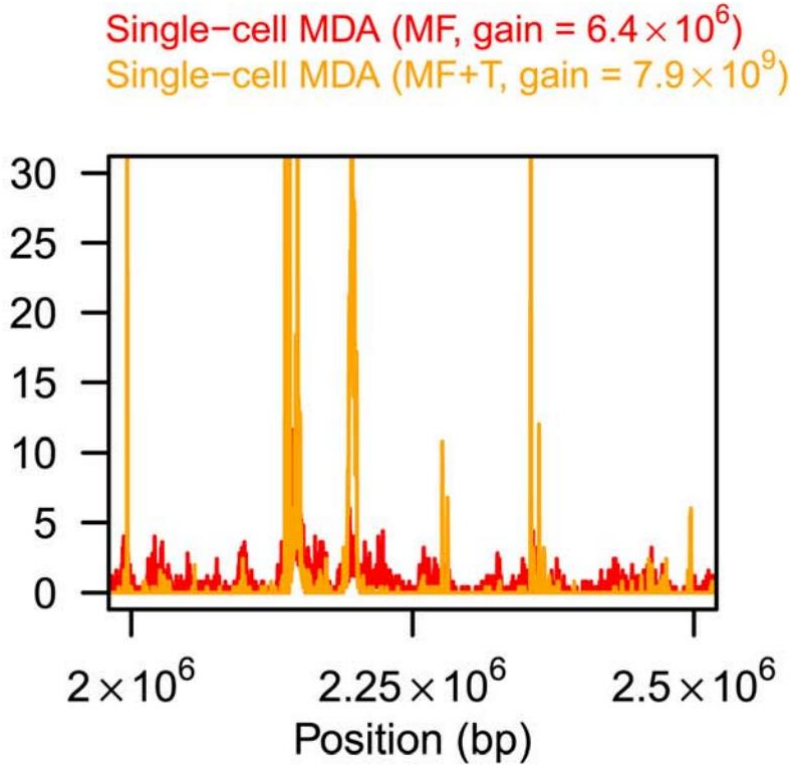

Figure 6.1 – Single Cell Sequencing Read Density vs Genome Position

(Modified from de Bourcy et. al., 2014)

Here, amplification bias is shown to be a function of reaction gain where genomes amplified to 10^6 (red) exhibit less pronounced sequence dependent amplification bias than reactions of 10^9 (orange).

In contrast to running additional rounds of WGA on already amplified product, studies such as the one done by Y. Hou et. al. in 2015 shows that if instead a second round of WGA is performed on a separate single cell with an identical genome as the first cell, they can improve overall coverage of target genome.[2]

This is because of one round of WGA on a single cell can be overlaid with sequences derived in subsequent rounds to obtain an overall improved genome coverage. However, thus far, this principle has only been demonstrated on separate aliquots of single cells, and not on a single cell alone. This is for the reason that in existing single cell WGA technologies, the gDNA resides in the same pool as the amplified material throughout WGA. Thus, after a single round of amplification, it is no longer possible to extract the original genomic DNA template for a second round of amplification; only WGA on amplified product is possible after the first WGA. On the other hand, our single cell micropillar array based microfluidic device and the GAMA process is able to retain the gDNA template in a fixed position within the microchannels while amplified product is washed away and collected in the output. Such a capability is especially critical for CTCs where mutations in the genome or epigenome may be the cause of cancer aggression.[7-8]

By entrapping single cells within our device, lysing it, extracting the genomic DNA, and performing multiple rounds of WGA in succession, we can use DNA sequencing or whole genome microarrays to determine the behavior of the bias on a single cell basis. In these experiments, if the same genes from the cell are present in each round of on-chip WGA, this indicates one of two possibilities: either that the bias is not random, or that we have lost a portion of the genome

during the extraction process. Additionally, simultaneous to genome coverage analysis, multiple rounds of WGA on single cells can be used to characterize the GAMA DNA extraction and WGA process. Numerous studies have shown that naked DNA slips off microfabricated pillar arrays when under hydrodynamic flow,[9-12] however, this has never been qualitatively observed within our device even after remaining suspended by flow for 52 hours. With the speed at which slipping occurs being a length dependent process, we can expect longer time scales of gDNA tethered within our device to show correlation with gDNA loss. Whether this phenomenon is occurring within our device can be individually determined by keeping a constant flow rate on extracted DNA and regularly collecting and sequencing the output port. Alternatively, this information can be analyzed simultaneous to multi-round WGA genome coverage analysis by checking for repeated gene early rounds to subsequent rounds. Likewise, whether or not template gDNA loss can occur through other means such as random physical breakage of the DNA can also be determined in a similar fashion.[13-14] Together, the GAMA process and the described micropillar array based microfluidic serves as a powerful single cell DNA processing and amplification platform.

6.3 REFERENCES

1. R. S. Lasken, T. B. Stockwell, BMC Biotech., 2007, 7(19), 1-11
2. Y. Hou, X. Shi, F. Li, L. Song, H. Wu, M. Dean, G. Li, S. Tsang, R. Jiang, X. Zhang, B. Li, G. Liu, N. Bedekar, N. Lu, G. Xie, H. Liang, L. Chang, T. Wang, J. Chen, Y. Li, X. Zhang, H. Yang, X. Xu, L. Wang, J. Wang, GigaScience, 2015, 4(37), 1-16
3. J. Hasmats, H. Green, C. Orear, P. Validire, M. Huss, M. Kaller, J. Lundeberg, PLOS One, 9(1), e84785, 1-10
4. M. Chen, P. Song, D. Zou, X. Hu, S. Zhao, S. Gao, F. Ling, PLOS One, 2014, 9(12), e114520, 1-12
5. C. F. de Bourcy, I. D. Vlaminck, J. N. Kanbar, J. Wang, C. Gawad, S. R. Quake, PloS One, 2014, 9(8), e105585
6. K. M. Ellegaard, L. Klasson, S. G. Andersson, PloS One, 2013, 8(11), e82319
7. A. Lyberopoulou, G. Aravantinos, E. P. Efstathopoulos, N. Nikiteas, P. Bouziotis, A. Isaakidou, A. Paplois, E. Marinos, M. Gazouli, PLOS One, 2015, 1-12

8. A. L. Hodgkinson, C. J. Morrow, Y. Li, R. L. Metcalf, D. G. Rothwell, F. Trapani, R. Polanski, D. J. Burt, K. L. Simpson, K. Morris, S. D. Pepper, D. Nonaka, A. Greystroke, P. Kelly, B. Bola, M. G. Krebs, J. Antonello, M. Ayub, S. Faulkner, L. Priest, L. Carter, C. Tate, C. J. Miller, F. Blackhall, G. Brady, C. Dive, Nature Medicine, 2014, 20, 897–903
9. S. W. P. Turner, M. Cabodi, H. G. Craighead, Physical Review Letters, 2002, 88(12), 128103, 1-4
10. J. D. P. Thomas, K. D. Dorfman, Biomicrofluidics, 2014, 8, 034115, 1-11
11. G. B. Salieb-Beugelaar, K. D. Dorfman, A. van den Berg, J. C. T. Eijkel, Lab Chip, 2009, 9, 2508-2523
12. S. M. Friedrich, H. C. Zec, T. H. Wang, Lab Chip, 16, 790-811
13. F. Farzaneh, S. Shall, A. P. Johnstone, Febs Letters, 1985, 189(1), 62-66
14. L. Davis, N. Maizels, PNAS, 2014, 111(10), E924–E932