MICROFLUIDICS FOR GENETIC AND EPIGENETIC ANALYSIS

SAI MA

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

Doctor of Philosophy
in
Biomedical Engineering

Chang Lu, Chair
Rafael V. Davalos
Adam R. Hall
Scott S. Verbridge
Hehuang Xie

April 28, 2017
Blacksburg, Virginia

Keywords: Microfluidics, single cell polymerase chain reaction, chromatin immunoprecipitation, bisulfite sequencing, next generation sequencing, cell electroporation, electrolysis

MICROFLUIDICS FOR GENETIC AND EPIGENETIC ANALYSIS

SAI MA

ABSTRACT

Microfluidics has revolutionized how molecular biology studies are conducted. It permits profiling of genomic and epigenomic features for a wide range of applications. Microfluidics has been proven to be highly complementary to NGS technology with its unique capabilities for handling small volumes of samples and providing platforms for automation, integration, and multiplexing. In this thesis, we focus on three projects (diffusion-based PCR, MID-RRBS, and SurfaceChIP-seq), which improved the sensitivities of conventional assays by coupling with microfluidic technology. MID-RRBS and SurfaceChIP-seq projects were designed to profiling genome-wide DNA methylation and histone modifications, respectively. These assays dramatically improved the sensitivities of conventional approaches over 1000 times without compromising genomic coverages. We applied these assays to examine the neuronal/glial nuclei isolated from mouse brain tissues. We successfully identified the distinctive epigenomic signatures from neurons and glia. Another focus of this thesis is applying electrical field to investigate the intracellular contents. We report two projects, drug delivery to encapsulated bacteria and mRNA extraction under ultra-high electrical field intensity. We envision rapid growth in these directions, driven by the needs for testing scarce primary cells samples from patients in the context of precision medicine.

MICROFLUIDICS FOR GENETIC AND EPIGENETIC ANALYSIS

SAI MA

GENERAL AUDIENCE ABSTRACT

Microfluidics is a technology that manipulates solution with extremely small volume. It is an emerging platform that has revolutionized how molecular biology studies are conducted. It permits profiling of genome wide DNA changes or DNA-related changes (e.g. epigenomics) for a wide range of applications. One of the major contribution of microfluidics is to improve the next generation sequencing (NGS) technologies with its unique capabilities for handling small volumes of samples and providing platforms for automation, integration, and multiplexing. In this thesis, we focus on three projects (diffusion-based PCR, MID-RRBS, and SurfaceChIP-seq), which improved the sensitivities of conventional assays by coupling with microfluidic technology. MID-RRBS and SurfaceChIP-seq projects were designed to profiling genome-wide DNA methylation and histone modifications, respectively. DNA methylation and histone modification have been proved to affect a lot of biological processes, such as disease development. These developed technologies would benefit the development of precision medicine (a medical model that proposes the customization of healthcare) and treatment to various diseases. We applied these technologies to study the epigenomic differences between several cell types in the mouse brain.

ACKNOWLEDGMENTS

The completion of this dissertation would not have been accomplished without the support from my teachers, friends, and family. They play critical roles in my life and help me become the person and researcher I am. I would like to express my gratitude to all of them because of their invaluable contributions to me.

First of all, I would like to thank my advisor and head of my committee, Dr. Chang Lu, for his patient guidance and support. He is more than my advisor, but also a sincere friend. Dr. Lu is the one that lead me into the wonderful field of microfluidics and epigenomics. His enthusiasm and critical thinking encouraged me to become a mature researcher. He took great care of me when I was sick, which I can never thank him enough. I greatly appreciate all his contribution of time, concerns, and financial support to make my PhD journal a rich experience. It is truly an honor to work with him. I hope I could become a successful, humorous, and intelligent person as Dr. Lu someday.

I would also like to thank the rest of my committee members, Dr. Rafael Davalos, Dr. Adam Hall, Dr. Scott Verbridge, and Dr. Hehuang Xie for their insightful suggestions and advices. Their supports during my prelim, dissertation, and job hunting have been priceless. Special thanks go to Dr. Hehuang Xie, and his graduate student Zhixiong Sun. Their knowledge in DNA methylation guided me through my experiments on on-chip bisulfite conversion. I would also like to thank Dr. Joseph Ecker, Dr. M. Margarita Behrens,

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

SYMBOL	*DESCRIPTION*
φ	Concentration
$\partial\varphi/\partial x$	Concentration gradient
5' UTR	5′ untranslated region
3' UTR	3′ untranslated region
AGP	(3-Aminopropyl)triethoxysilane /Glutaraldehyde/Protein A
APTES	(3-Aminopropyl)triethoxysilane
a. u.	Arbitrary units
AUC	Area under the curve
CDS	Coding sequence
CFU	Colony forming units
CGI	CpG island
ChIP-seq	Chromatin immunoprecipitation coupled with deep sequencing
CN	Cell number
CPP	Cell penetrating peptides
CpG site	5'-C-phosphate-G-3
Ct	Cycle threshold
D	Diffusivity or diffusion coefficient
DAQ card	Data acquisition card
dATP	Deoxyadenosine triphosphate
DC	Direct current
dCTP	Deoxycytidine triphosphate

DE	Differentially expressed
dGTP	Deoxyguanosine triphosphate
DI water	Deionized water
DIC	Differential interference contrast
DMEM	Dulbecco's modified eagle medium
DMR	Differentially methylated regions
dNTP	Nucleoside triphosphate
dsDNA	Double stranded DNA
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
ENCODE	Encyclopedia of DNA Elements
FACS	Fluorescence-activated cell sorting
FBS	Fetal bovine serum
FDR	False discovery rate
FPKM	Fragments per kilobase of transcript per million mapped reads
GFP	Green fluorescent protein
GO	Gene ontology
H3K27ac	Histone 3 lysine 27 acetylation
H3K27me3	Histone 3 lysine 27 tri-methylation
H3K4me3	Histone 3 lysine 4 tri-methylation
HCPs	High CpG-density Promoters
ICPs	Intermediate CpG-density Promoters
J	Diffusion flux
$(KFF)_3K$-O-PNA	H-KFFKFFKFFK-o-cataagacggt-NH_2
LB	Lysogeny broth

LCPs	Low CpG-density Promoters
LinDA	Single-tube linear DNA amplification
LPS	Lipopolysaccharide
MID-RRBS	MIcrofluidic Diffusion-based RRBS
MNase	Micrococcal Nuclease
MOI	Multiplicity of Infection
MOWChIP	Microfluidic oscillatory washing–based ChIP
mRNA-seq	mRNA sequencing
M. smegmatis	Mycobacterium smegmatis
NGS	Next generation sequencing
NTCs	No-template controls
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PDMS	Polydimethylsiloxane
PEG	Polyethylene glycol
PFA	Perfluoroalkoxyalkane
PIC	Protease inhibitor cocktail
PMSF	Phenylmethylsulfonyl fluoride
PNA	Peptide nucleic acid
PS	Penicillin streptomycin
qPCR	Quantitative PCR
qRT-PCR	Quantitative reverse transcription PCR
RFU	Relative fluorescence units
*rpo*D	RNA polymerase sigma 70 (sigma D) factor gene
RRBS	Reduced representative bisulfite sequencing

SE	Super enhancer
SDS	Sodium dodecyl sulfate
ssDNA	Single stranded DNA
tRNA	Transfer RNA
TSA	Tryptic soy agar
TSB	Tryptic soy broth
TSS	Transcription start site
TTS	Transcription termination site
WGBS	Whole genome bisulfite sequencing

CHAPTER 1 – Overview

The scope of this research is to develop and optimize microfluidic protocols for novel molecular biotechnologies. The application of conventional techniques to microfluidic chip format provides unique advantages of improved sensitivity, speed, reproducibility, and reduced cost. This dissertation presents the development of five different microfluidic assays that capable of conducting genetic or epigenetic studies.

Cellular heterogeneity is evident in functionality [1, 2] and gene expression measurements (RNA-seq) within apparently homogeneous samples [3-5]. To analyze the genetic contents from cells, polymerase chain reaction (PCR) is the primary tool. One of the major limitations of PCR is the sensitivity and integration with upstream technologies. We applied a diffusion-based technology to perform PCR from intact cells (Chapter 3). This technology integrated multiple steps in a single device, including cell trapping, cell lysis, surfactant removal, and DNA amplification. We achieved sensitivity down to single cell level.

The gene expression and cellular behavior are not only affected by the DNA sequence, but also the epigenetic activities, such as DNA methylation and histone modification. The epigenetic activity shows great impact on the cell differentiation, self-renewal and disease development; however, the mechanism is still barely to be understood despite all the efforts made over the years. A better understanding of epigenetic regulation will benefit disease diagnosis and developing personalized medicine and therapeutics.

DNA methylation is one of the major epigenetic mechanisms. The gold standard to profile methylome is using bisulfite conversion coupled with sequencing. The bisulfite conversion is a chemical reaction, which converts unmethylated cytosines to thymines while keep methylated cytosines unchanged. The major drawback of bisulfite conversion is the significant DNA loss up to 96%, which led to large amount of starting materials (microgram level DNA). We applied a diffusion-based technology to conduct bisulfite conversion on microfluidic chip, which preserved 30% of the DNA and improved the sensitivity to sub-1 ng (Chapter 4).

The epigenetic regulation involves the interaction between DNA sequence and specific proteins (such as histone proteins). The histone modification is another of the major epigenetic mechanisms. To identify the genome-wide DNA/protein interaction, chromatin immunoprecipitation coupled with next generation sequencing (ChIP-seq) has become the most sensitive method. ChIP-seq involves gathering large amount of cells (10^6 ~10^7 cells), cell lysis, chromatin fragmentation by enzyme digestion or sonication, immunoprecipitation to capture target DNA/protein complex, proteinase K digestion to release DNA, DNA library construction/amplification and next generation sequencing for identifying enriched sequences. There are two major limitations of the conventional ChIP-seq protocols: *1)* requirement of large number of cells (>10^7 cells) which limits application for scarce samples, *2)* tedious process with long hands-on time, which compromises the reproducibility and throughput of the assay. We demonstrated a microfluidic technology, referred to as Surface Immobilized Antibody for Chromatin Immunoprecipitation coupled with next generation sequencing (SurfaceChIP-seq) for low-input (30-100 cells) profiling of multiple histone modification landscapes in parallel (Chapter 5).

The thesis is divided in two parts, based on the absence (part I) or presence (part II) of electrical methods in the system. Part I mainly discusses microfluidic technologies for genetic and epigenetic analysis. The diffusion-based approaches were developed for conducting PCR (Chapter 3) or bisulfite conversion on-chip (Chapter 4). We also developed a device with surface immobilized antibody for low cell number chromatin immunoprecipitation (Chapter 5). Part II focuses on ultra-high field intensity electrolysis of *Mycobacteria smegmatis* (Chapter 6) drug delivery using electroporation (Chapter 7). The focus of each chapter is outlined briefly:

Part I

CHAPTER 3 – Diffusion-based Microfluidic PCR for "One-pot" Analysis of Cells

Genetic analysis starting with cell samples often requires multi-step processing including cell lysis, DNA isolation/purification, and polymerase chain reaction (PCR) based assays. When conducted on a microfluidic platform, the compatibility among various steps often demands complicated procedure and complex device structure. Here we present a microfluidic device that permits "one-pot" strategy for multi-step PCR analysis starting from cells. Taking advantage of the diffusivity difference, we replace the smaller molecules in the reaction chamber by diffusion while retaining DNA molecules inside. This simple scheme effectively removes reagents from the previous step to avoid interference and thus permits multi-step processing in the same reaction chamber. Our approach shows high efficiency for PCR and potential for a wide range of genetic analysis including assays based on single cells.

CHAPTER 4 – A Microfluidic Device for Low-input Methylomic Analysis Using Sub-1 ng Starting DNA

Methylomic analyses require substantial amounts of DNA and this restriction hinders studies involving scarce animal and patient samples with direct biomedical relevance. Here we describe microfluidic diffusion-based reduced representative bisulfite sequencing (MID-RRBS) that permits methylomic profiling with 0.3 ng starting DNA. Using this technology, we examined genome-wide DNA methylation in neurons and glia isolated from mouse cerebellum and established cell type-specific methylomes. Our data revealed that genome-wide methylation in general and methylation in low CpG-density promoter regions in particular showed distinct patterns for neurons and glia. Furthermore, we found very little overlap between differentially methylated region-associated genes and differentially expressed genes when the two cell types were compared.

CHAPTER 5 – Paralleled Microfluidic Device for High-throughput and Low-input Histone Modification Analysis Based on Surface ChIP (surfaceChIP-seq)

The sensitivity and throughput of chromatin immunoprecipitation (ChIP) assays hinders the epigenomic study for scarce and clinical samples. We developed a microfluidic approach for automated and paralleled ChIP-seq assay using as few as 30-100 cells. Comparing to conventional approaches, modification-specific antibodies were immobilized on the glass surface in the microfluidic channels instead of beads (SurfaceChIP-seq). This approach simplifies the chip structure and allows up to simultaneously profiling eight assays. Using this technology, we revealed the epigenomic

signatures in neuron and glia isolated from mouse cerebellum, suggesting that the distinctive histone modification landscapes exist in brains.

Part II

CHAPTER 6 – RNA extraction from a mycobacterium under ultrahigh electric field intensity in a microfluidic device

Studies of transcriptomes are critical for understanding gene expression. Release of RNA molecules from cells is typically the first step for transcriptomic analysis. Effective cell lysis approaches that completely release intracellular materials are in high demand especially for cells that are structurally robust. In this report, we demonstrate a microfluidic electric lysis device that is effective for mRNA extraction from mycobacteria that have hydrophobic and waxy cell walls. We used packed bed of microscale silica beads to filter *M. smegmatis* out of the suspension. 4000-8000 V/cm field intensity was used to lyse *M. smegmatis* with long pulses (i.e. up to 30 pulses that were 5 s long each). Our qRT-PCR results showed that our method yielded a factor of 10-20 higher extraction efficiency than the current state-of-the-art method (bead beating). We conclude that our electric lysis technique is an effective approach for mRNA release from hard-to-lyse cells and highly compatible with microfluidic molecular assays.

CHAPTER 7 – Electroporation-based delivery of cell-penetrating peptide conjugates of peptide nucleic acids for antisense inhibition of intracellular bacteria

Cell penetrating peptides (CPPs) have been used for a myriad of cellular delivery applications and were recently explored for delivery of antisense agents such as peptide nucleic acids (PNAs) for bacterial inhibition. Although these molecular systems (i.e. CPP-PNAs) have shown ability to inhibit growth of bacterial cultures in vitro, they show limited effectiveness in killing encapsulated intracellular bacteria in mammalian cells such as macrophages, presumably due to difficulty involved in the endosomal escape of the reagents. In this report, we show that electroporation delivery dramatically increases the bioavailability of CPP-PNAs to kill Salmonella *enterica* serovar Typhimurium LT2 inside macrophages. Electroporation delivers the molecules without involving endocytosis and greatly increases the antisense effect. The decrease in the average number of Salmonella per macrophage under a 1200 V/cm and 5 ms pulse was a factor of 9 higher than that without electroporation (in an experiment with a multiplicity of infection of 2:1). Our results suggest that electroporation is an effective approach for a wide range of applications involving CPP-based delivery. The microfluidic format will allow convenient functional screening and testing of PNA-based reagents for antisense applications.

CHAPTER 2 – Microfluidics for Epigenomic Analysis

2.1 Introduction

Next-generation sequencing (NGS) is transforming the understanding of biology. Since the finish of the human genome project in 2003, the cost of sequencing has significantly decreased which makes it available to more researchers. Unfortunately, advances in sequencing technologies are not always accompanied by the development of new sample processing procedures. Genomics and transcriptomics (the study of the complete set of DNA or RNA of cells at a specific stage) have already achieved single cell sensitivity using conventional tube-based approaches [6-11]. Tube-based approaches have hit their limit in terms of amplification bias and throughput (only up to tens of samples each batch). Such methods are not particularly suitable for investigating the heterogeneity among single cells due to manual handling errors and vast number of targeted cells. Epigenomics is the study of heritable modifications on DNA or histone without changing DNA sequences. It is an emerging field that usually requires large amount of starting material for the assays (e.g. ChIP-seq, MeDIP-seq and Bisulfite-seq). Benchtop versions usually do not provide an efficient way to improve the sensitivity of these assays.

Microfluidics, which allows manipulation of solution with extremely small volume, has been widely used for monitoring individual cells [7, 12, 13] and manipulating microenvironments [14, 15]. Utilizing parallel structures, microfluidic devices are capable of processing hundreds of samples simultaneously within their tiny chambers. The miniaturized structures improve throughput, reduce reagent amount, and the amount of

starting materials. Microfluidic devices are becoming an important tool to improve the performance of sequencing related molecular technologies [12, 16-19].

2.2 Epigenomics

Epigenomics is the freshly developed study of epigenetic modification of a cell which is analogous to genomics and proteomics [20]. Epigenetics refers to the phenomena that gene expression changes without changing DNA sequencing, in another word, a change in phenotype without a change in genotype. Epigenetic change is a regular and natural, but also involves age, environment and diseases. Epigenetic change can determine common manner, such as cell differentiation or more damaging effects, such as cancers. Epigenetic modifications involve the reversible and heritable modifications on DNA or histone without changing DNA sequences. The most commonly studied epigenetic modifications are histone modification, DNA methylation, and non-coding RNA (ncRNA). These tags change the chromatin structure, DNA accessibility and regulation of gene expression. It has been proved that these modifications play critical roles in a lot of biological processes, such as brain development [21] and tumor formation [22]. Unlike genomic analysis, most of the epigenomic assays require a large amount of starting materials. For example, 1 million cells are required for histone modification profiling, and microgram levels of DNA is needed for DNA methylation analysis. To overcome this drawback, several microfluidic based approaches [23-31] have been developed and dramatically improved the sensitivity and throughput of conventional assays. It can be seen that utilizing microfluidic technologies for epigenomic profiling is a fast growing field.

DNA methylation

In eukaryotic DNA, it is known that the methylation occurs at the cytosine which are converted to 5-methylcytosine (5-mC). The DNA methylation is believed to repress the gene expression by blocking the promotors where the transcription factors are supposed to bind. Extensive studies have demonstrated that DNA methylation plays a major role in many biological phenomena, such as cellular proliferation, differentiation, and various diseases [32, 33]. Some studies have also indicated that there is connection between histone modification and DNA methylation at certain genomic positions [34, 35]. It has also been demonstrated that DNA methylation is implicated in brain development underlying learning and memory [21].

Chromatin structure and histone modification

Within the nucleus of eukaryotic cells, DNA is highly condensed and wrapped around nuclear proteins (histone). The complex of DNA and histone is called chromatin. The basic chromatin unit, the nucleosome is formed by a ~147 bp DNA chain wrapped around an 8 histone subunits. The chromatin can be fragmented by sonication or enzyme digestion (e.g. micrococcal nuclease). Under optimized condition, mono-nucleosomes can be obtained from chromatins. There are five types of histone subunits: H1, H2A, H2B, H3 and H4. H1 stabilizes the chromatin structures by binding to the linker region. H2A, H2B, H3 and H4, which are called core histones, form nucleosome by two copies of each. Nucleosomes connect to each other by a short region of DNA and form "beads on a string" fiber which is the primary structure of DNA packaging. Multiple "beads on a string" fibers are packed together to form 30 nm solenoid which extends to chromosome.

The core histones are mostly globular except for their unstructured "tails". The most important feature of histone, especially the "tails", is the large amount and types of residues they process. There are more than eight modifications have been found on histone. The most commonly studied modifications are acetylation, methylation and phosphorylation which are named after the different functional groups attached. Histones contains more than 60 different residues that can be modified. Besides, methylation that happens on lysines or arginines may be one of three different forms: mono-, di-, or trimethyl for lysines and mono- or di- (asymmetric or symmetric) for arginines. For example, H3K4me3 means trimethylation of the fourth amino acid residue from the N-terminus of histone H3. The status of all these modification depends on the cell type, status and environment. To monitor the global incidence of histone modification, modification-specific antibody is used to capture histone associated DNA coupled with next gen sequencing (ChIP-seq).

State of the art of ChIP-seq

ChIP-seq involves cell lysis, chromatin fragmentation, immunoprecipitation of the protein of the interest, protein digestion, DNA enrichment, and sequencing. Conventional ChIP-seq assay usually requires about 10 million cell which is impractical for precious samples. Various efforts have been made to improve the conventional protocol over the last few years, including CChIP [36], Nano-ChIP [37, 38], LinDA [39], iChIP [40], ULI-NChIP [41], WGA-ChIP-seq [42] and FARP-ChIP-seq [43] and Drop-ChIP (single-cell ChIP-seq) [31]. By utilizing microfluidic devices, our lab developed MOWChIP-seq to detect histone marks (H3K4me3 and H3K27Ac) from as little as 100 cells which is described in section 2.5.2. In the section 3.3, a novel glass surface antibody

immobilization approach is described which performs similar to MOWChIP-seq but is suitable for single cell ChIP assay.

"Carrier ChIP", (CChIP [36]), represents a category of ChIP protocol using carrier chromatin or mRNA to reduce the non-specific adsorption of magnetic beads. The original CChIP protocol detected H4K16ac, H3K4me3, H3K9me2 and H3K4me1 histone marks from 100 to 1000 mouse ES cells. Using carrier chromatin is an effective way to improve the sensitivity of ChIP-qPCR; however, the carrier chromatins will significantly affect the sequencing depth when used in ChIP-seq assay. Using mRNA carrier which is digested later has been shown to increase recovery of ChIP DNA and allowed transcription factor ChIP-Seq using 10,000 cells [44]. RP-ChIP-seq and FARP-ChIP-seq protocols [43] was published in November, 2015 using similar carrier approach. In RP-ChIP-seq, yeast chromatins were used as carrier for immunoprecipitation of mouse chromatin, about 1~2 million reads are finally mapped to mouse genome which was not sufficient compare to literature [45]. To improve the sequencing depth of RP-ChIP-seq, Favored Amplification RP-ChIP-Seq (FARP-ChIP-seq) was developed. A 210-bp biotin-DNA that does not map to Drosophila, mouse, or human genomes was designed as carrier. The biotin-DNA and the chromatin of interest were recovered by streptavidin beads and antibody-coupled protein G beads, respectively. To inhibit amplification of biotin-DNA during library construction, a PCR amplification blocker oligo complementary to the biotin-DNA was designed to contain modifications at both ends. The blocker oligo reduced the amplification of the biotin-DNA by over 99%. RP-ChIP-seq examined histone mark H3K4me3 using 500 ES cells.

Nano-ChIP [37, 38] was developed by Bradley E Bernstein's group to examine histone modification using 5,000 to 10,000 cells. Single-tube linear DNA amplification (LinDA [39]) was developed to profile the histone mark (H3K4me3) using 10,000 cells, oestrogen receptor-α (ER-α) binding using 5000 cells and transcription factor from as little as 30 pg DNA. Both nano-ChIP and LinDA explored their own DNA amplification strategy. The primer with hairpin structure was used in Nano-ChIP to minimize primer self-annealing. The primers containing restriction sites were digested by BciVI and used directly for sequencing adaptor ligation. In LinDA, ChIP DNA was ligated with poly T and and *in vitro* transcripted to RNA. The RNA was then reverse transcribed and amplified using the T7 promoter-BpmI-oligo(dA)15 primer. Linda was based on an optimized T7 phage RNA polymerase linear amplification protocol that reduces amplification bias due to GC content. In WGA-ChIP-seq [42], purified DNA was amplified by WGA4 kit, a single cell whole genome amplification kit from Sigma-Aldrich. The adaptors were removed by BpmI restriction enzyme and used for library construction.

Indexing-first ChIP (iChIP [40]) was recently developed to index each sample and pool together before ChIP. Magnetic beads coated with anti-H3 hisone antibody was used to capture chromatins. The chromatins (~500 cells) were end-repaired, A-tailed, and ligated with indices. The indexed chromains were then pooled and subjected to second round of immunoprecipitation with specific antibody (H3K4me3, H3K27ac et al.). The ChIP DNA from pooled samples was then sequenced and the data were demultiplexed based on indices to yield a sensitivity of 500 cells per individual sample.

ULI-NChIP [41] was developed to work with native-ChIP (i.e. epigenetic mark does not crosslinked with DNA) that the chromatins were sheared by MNase digestion. There were able to detect histone mark (H3K27me3, H3K9me3 and H3K4me3) from 1000 cells.

2.5 Next Generation Sequencing

Nucleic acid sequencing refers to methods for determining the axact order of nucleotides in DNA or RNA. Since the Human Genome Project (HGP) which took 13 years and cost $3 billion completed in 2003, the use of sequencing has exponentially increased [46]. Sanger sequencing which is known as fist generation sequencing, was used to accomplish HGP. Since the Sanger sequencing was developed in 1973, the throughput of sequencing has been significantly improved while the cost has been significantly reduced. The second generation sequencing (next generation sequencing, NGS) is able to perform massively sequencing which means millions DNA fragments can be sequenced simultaneously. It has been widely used for all kinds of applications, including whole-exome sequencing, RNA-seq, whole genome sequencing, methylation sequencing, and ChIP-seq.

There are several NGS platform has been commercialized. 454 pyrosequencing (Roche Applied Science), Solexa technology (Illumina Genome Analyzer), SOLiD platform (Applied Biosystems), ion torrent semiconductor sequencing and Hi-seq (Illumina Hi-seq 2500) are the most popular platforms. Among them, Illunima Hi-seq is the most commonly used platform providing >200 million reads per lane for less than $1000. Illunima sequencing is based on cluster generation and sequencing by synthesis (SBS) technology. To be able to be sequenced by Illumina sequencing, target DNA

usually needs to be end-repaired and adaptor ligated on both ends. After denaturation, single strand DNA fragments containing adaptors are immobilized to flow cell surface with complimentary adaptors. Each immobilized fragment creates a bridge structure by hybridizing its free end with another complimentary adaptor. To achieve sufficient fluorescence signal for identify dNTPs, immobilized fragments are amplified through PCR, which is known as bridge amplification. The sequencing templates were amplified to generate up to 1000 identical copies per cluster by solid phase amplification. The density of cluster can achieve more than 10 million clusters per square centimeter. Four fluorescently-labelled nucleotides (dNTPs) are used to sequence clusters. During each sequencing cycles, one dNTPs is incorporated on each DNA copy. The dNTP is identified by scanning the fluorescence signals during enzyme cleavage.

The principle of NGS is very similar to capillary electrophoresis (CE) sequencing. DNA synthesis is catalyzed by polymerase to add fluorescently labelled dNTPs onto the DNA templates during a series of cycles. At the end of each cycle, the fluorescent signals are analyzed to identify the added nucleotides. NGS allows processing millions of fragments in parallel, which significantly improves throughput and decreases sequencing costs. There are five major steps to prepare a sequencing sample/library: DNA fragmentation, end-repair/A-tailing, adapter ligation, amplification, and quality control/sample pooling (**Fig. 1**).

DNA fragmentation

End repair

A-tailing

Adapter ligation

PCR

Sequencing
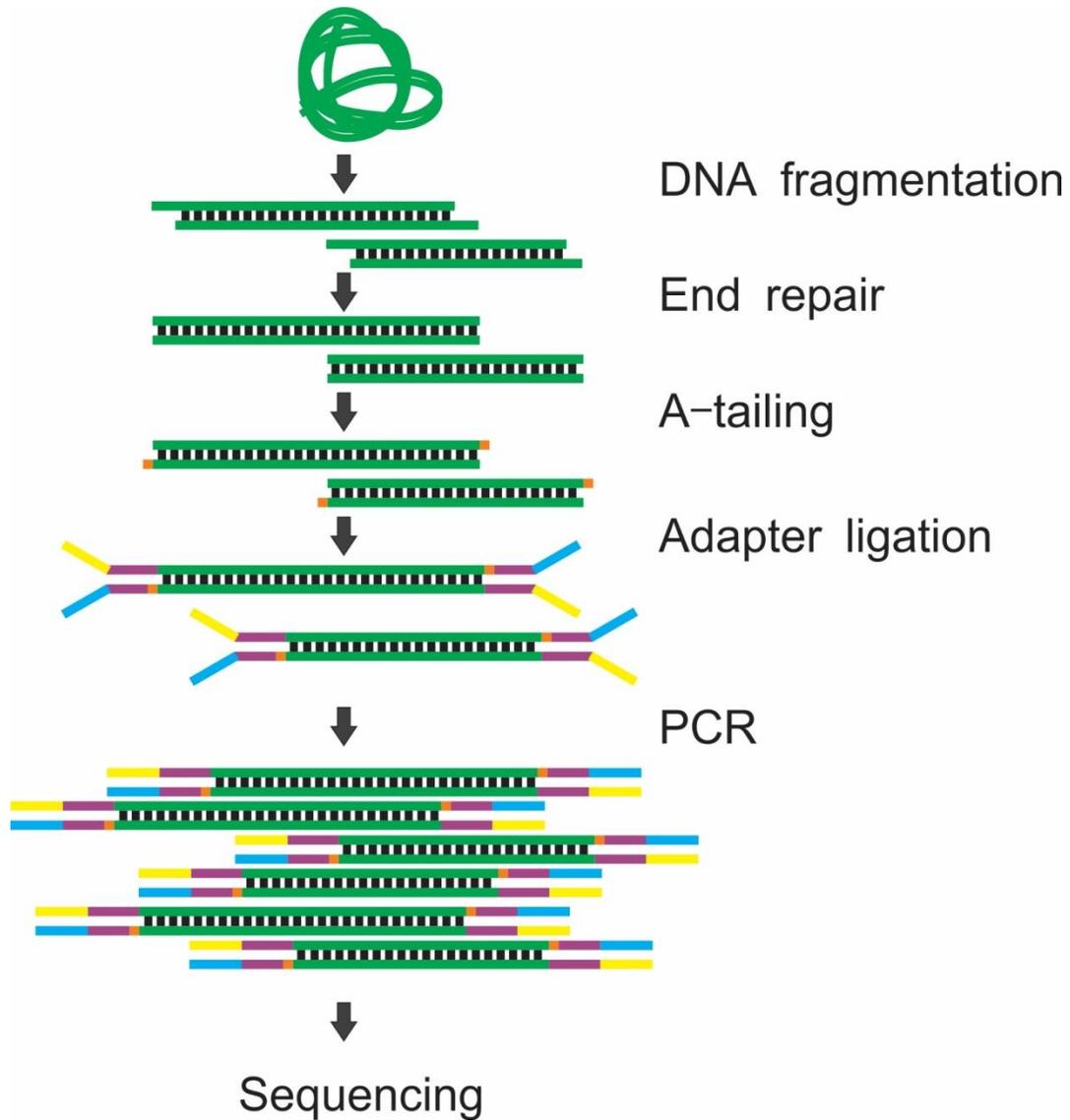
Fig. 2.1 The NGS library preparation procedure.

The starting material for library preparation needs to be fragmented to ~100-500 bp to be properly amplified in the flow cell. The fragmentation is typically performed by sonication, enzyme digestion, or tagmentation. Next, a pair of sequencing adapters are ligated to both ends of the DNA. The ligated DNA will be able to attach to the flow cell by hybridization during sequencing. To maximize the ligation efficiency, the DNA fragments

are usually subjected to end-repair and A tailing (generate adenine overhang to the 3' end of DNA) before ligation. Depending on the amount of starting materials, the ligated products may be insufficient for sequencing (≥ 2 nM). Polymerase chain reaction (PCR) is the most commonly used approach to amplify the ligated products and generate enough material for sequencing. The fragment sizes of amplified products are checked by gel electrophoresis or Bioanalyzer to ensure they are properly ligated. The libraries are then pooled to the desired concentration (2-10 nM) and are ready for sequencing. Depending on the applications, the library preparation procedure may be quite different. We will discuss various library preparation procedures for the three major fields, genomics, transcriptomics and epigenomics.

The final step in the NGS pipeline is genome mapping. The obtained reads need to be aligned to a known reference genome, such as hg19, mm8. Depending on the reads quality, the raw reads might be trimmed to reduce bias and improve alignment rate. Usually, uniquely mapped reads with length longer than 20-30 bp will be used for following analysis. After alignment, the DNA sequence of each tag is replaced with the genomic coordinates.

2.3 Sequencing-related Microfluidic Technologies for Epigenomics Analysis

In recent years, the microfluidic applications for epigenomic analysis are rapidly growing. The mechanisms of these applications can be classified into three categories, bisulfite-based approaches, affinity-based approaches, and digestion-based approaches.

Bisulfite-based approaches

To look into the genome wide DNA methylation profile, the most established technology is a bisulfite conversion based analysis coupled with next generation sequencing, which is regarded as the gold standard for DNA methylation analysis. Bisulfite conversion enables methylation analysis at single base resolution [47-49]. It is based on the difference in reactivity of sodium bisulfite with unmethylated and methylated DNA. Sodium bisulfite converts the unmethylated cytosines to uracils while the methylated cytosines remain the same. After conversion, the original DNA sequences are distinguished by quantitative PCR or sequencing. For bench approaches, the sensitivity of bisulfite sequencing has been improved to single cell level [50, 51]. It is also coupled with other sequencing technology (RNA-seq) [52, 53] to obtain more information from the same cell.

Several efforts have been made to perform bisulfite conversion on a microfluidic chip. Shin et al. designed a droplet based platform for bisulfite conversion [54]. The DNA is bound to the surface of magnetic silica beads under low pH and high chaotropic salt concentration. The DNA can be released by reversing these conditions. By moving the DNA/bead complex, the DNA is transferred into different buffers for conversion and

cleanup. For comparison, Yoon et al.[55] used a glass surface as the substrate to adsorb DNA instead of silica beads. The denatured DNA is first mixed with bisulfite cocktail and converted on chip. The converted DNA is mixed with guanidine hydrochloride and adsorbed to the glass surface by electrostatic interaction. In all these cases, the DNA is detected by PCR for analyzing the methylation status. Some efforts have also been made to improve the accuracy and throughput of PCR using microfluidic devices [56-59]. Although the on-chip conversion methods could be used for genome-wide DNA methylation profiling, no such demonstration has been published yet. (cite my paper after publishing)

Affinity-based approaches

i. ChIP

In the eukaryotic cell, DNA wraps around globular histone cores and forms nucleosomes. The nucleosomes pack together and form chromatin. The histone tails can be covalently modified, such as acetylation, methylation, phosphorylation, and ubiquitination. Acetylation and methylation are the most commonly studied modifications. Chromatin immunoprecipitation (ChIP) is the primary technology used to examine histone modifications [60]. It can also be used to determine if a specific protein (i.e. transcription factors) interacts with certain genomic region (e.g. enchancer and promotor). ChIP is divided into two categories, depending on the method used to process the chromatin, XChIP and NChIP. For NChIP (native ChIP), the chromatin is not cross-linked so that it is suitable for mapping histone marks or transcription factors that strongly bind to DNA. Usually the chromatin is fragmented by micrococcal nuclease (MNase) digestion to yield a ladder of DNA fragments corresponding to the size of multiple nucleosome cores plus the linker (~170 bp). For XChIP (cross-linked ChIP), the chromatin is stabilized by

formaldehyde, which makes it difficult for MNase to access to chromatin. Instead of enzyme digestion, sonication is typically used for XChIP. The fragmented chromatin (by sonication or enzyme digestion) is selectively targeted by specific antibodies that are immobilized to magnetic or agarose beads' surface. The selected chromatin/bead complexes are washed, and then the DNA is eluted and purified. The ChIPed DNA can be detected by quantitative PCR (ChIP-qPCR) for examining a few loci, NGS (ChIP-seq), or microarray (ChIP-chip) for genome wide profiling. Using a conventional assay, genome wide profiling (~10 million cells) usually requires more dedicated optimization and at least one order of magnitude more starting materials than ChIP-qPCR (~1 million cells).

ii. ChIP-qPCR

The recent advances of microfluidic ChIP-qPCR were reviewed by Matsuoka et al [61]. The first applications of ChIP on a microfluidic platform were independently developed by Oh et al. [23] and Wu et al. [24]. Oh et al. designed a flat chamber (bead reservoir) to hold micro-agarose beads [23]. The bead reservoir is connected to the dispersion channels via short and shallow channels, which stops microagarose beads while allowing solution to flow through. This chip design does not involve the use of complicated micro-valves. It requires 2.5×10^6 cells as starting material, which is similar to the performance of the conventional assay. Wu et al. designed AutoChIP [24] and HTChIP [25] for automated high throughtput ChIP-qPCR analysis using sheared chromatin corresponding to 1000-2000 cells for each sample. Chromatin is actively mixed with antibody-beads complexes in a circulating peristalic mixer. The chromatin-antibody-bead complexes are then stacked to a column by micromechanical valves and washed by buffers. They were able to perform ChIP assay from 4 and 16 samples simutanusly.

Geng et al. from our group developed the most sensitive on-chip ChIP-qPCR assay [26]. This protocol utilizes N-ChIP (native chip) coupled with MNase digestion. Intact cells are directly loaded on chip instead of sonicated chromatin. The cells are lysed and digested by MNase. The fragmented chromatin is then forced to flow through a pre-packed antibody-bead bed in a connected micro-chamber. Oscillatory washing is utilized to yield high quality ChIP DNA. This technique allowed sensitive qPCR with as few as 50 cells.

Cao et al. from our group integrated sonication into a microfluidic chip for histone modification and DNA methylation profiling [27]. The chromatin or DNA is sheared by transducer, which is attached to the glass bottom of the chip. Compared to Geng's protocol [26], Cao's approach worked with cross-linked cells (X-ChIP), which may extend the application to profile transcription factors. These two approaches are complimentary to each other and make the on-chip ChIP assay suitable for both native chromatins and cross-linked chromatins. Cao's work reaches similar sensitivity (~100 cells) as Geng's work with significantly improved signal-to-noise ratio (fold enrichment).

iii. ChIP-seq

Even though, on-chip ChIP-qPCR assay achieved very good sensitivity, it can only be used to examine a few loci for each sample. To extend to the analysis to genome wide profiling from a few loci, several efforts have been made by taking advantage of microfluidics.

Jiang et al. developed a microfluidic chip for profiling H3K4me3 landscape by ChIP-seq [28]. The chip shared similar structure as the chip developed by Wu et al. [24, 25]. The

chromatin and beads are circulated in a dead-end flow channel for immunoprecipitation. The chromatin-antibody-bead complex is trapped to form a column and washed. This protocol was able to detect histone mark from 1000 mouse early embryonic cells.

Recently, our group developed the microfluidic oscillatory washing–based ChIP-seq (MOWChIP-seq) [29] protocol for histone marks (H3K4me3 and H3K27ac) from as little as 100 cells. In the MOWChIP-seq, chromatin is flowed through a packed bed of antibody coated beads. The packed bed leads to high-efficiency adsorption and increases the chance of nonspecific adsorption and physical trapping. The chromatin/bead complexes are washed by oscillatory washing in two buffers which effectively removed non-specifically absorbed chromatins.

The Drop-ChIP protocol, developed by Rotem et al., collected ChIP-seq data at single cell level [31]. Their strategy was to extract and digest chromatin of single cells in individual droplets. The fragmented chromatin is labeled by unique DNA barcodes so that each cell can be distinguished after sequencing. Chromatin collected from about 100 cells is pooled together for conventional ChIP assay. Carrier chromatin (i.e. chromatin extracted from other species) is used to minimize non-specific adsorption. The carrier chromatin unfortunately results in huge amount of junk reads. They only collected 7 million useful reads (700,000 unique reads) from 320 million reads in total. On average ~1000 unique reads are obtained for each single cell. Typically, for conventional ChIP-seq of bulk sample, 10-20 million reads are preferred for the same histone mark. Even though they provided an elegant solution to investigate heterogeneity among thousands of single cells, it provided too little information compared to the other type of single cell sequencing technology, such as RNA-seq and Bisulfite-seq.

iv. MeDIP and MBD

The affinity-based approaches, including Methylated DNA Immunoprecipitation sequencing (MeDIP-seq) [62, 63] and Methylated DNA Binding Domain sequencing (MBD-seq) [64] are also used for profiling DNA methylation status. These technologies require much more starting materials (1-300 ng).

Microfluidic devices can also be used to enrich methylated DNA. Methyl-binding domain (MBD) protein [65, 66] or 5-methylcytidine (5-mC) antibody [27] which specifically targets methylated DNA is immobilized on magnetic beads or the surface of the microfluidic chamber. When the DNA mixture contact MBD protein or 5-mC antibody, the methylated DNA is captured. The non-captured DNA is washed away and the captured DNA is then eluted for qPCR analysis. These methods have been demonstrated for examining the methylation status of specific loci [67]. They have not been applied for genome wide analysis.

v. Transcription factor binding affinity

Transcription factors (TFs) are proteins that bind to specific DNA sequences so that they control the transcription rate of DNA. The most common way to profile transcription factor binding sites is using ChIP-seq, which is similar to histone modification detection. The TFs that bind to DNA are fixed at certain time point. The DNA is then fragmented by sonication and the TF/DNA complex is specifically selected by the antibody. Because there is a probability that the TFs actually bind to DNA, ChIP-seq for TFs requires much more starting material than that for histone modification. The binding strength between TFs and DNA is also usually much weaker compared to that between histone and DNA.

Due to the above reasons, ChIP-seq for TFs has not been achieved in microfluidic platform.

Maerkl et al. developed an alternative to systematically study the binding affinity of TFs, especially low-affinity interactions [68-73]. The device contains 2400 units and each unit is controlled by three micromechanical valves and a "button" membrane. The button membrane is used for surface derivation and control molecular interaction. The chip surface is locally derived with antibody to capture target DNA and TFs. It provides a way for large-scale quantitative protein-DNA interaction measurement, which can be used to verify and predict the *in vivo* function of TFs.

Digestion-based approaches

The way that the DNA is packaged into chromatin is critical for gene regulation. Several methods have been developed for analyzing chromatin conformation, accessibility, and nucleosome positioning [74]. To evaluate the chromatin accessibility, specific enzymes (MNase, DNase and transposase) are used to digest chromatin. Combined with NGS, these methods are used for profiling genome-wide chromatin status. These methods include chromosome conformation capture (Hi-C) [75], assay for transposase-accessible chromatin (ATAC-seq) [76], DNase-seq [77], MNase-seq [78], and formaldehyde-assisted isolation of regulatory elements with sequencing (FAIRE-seq) [79]. Hi-C is used to identify the 3D conformation of chromatin. MNase-seq identifies the nucleosome positioning by digesting chromatin with MNase. In DNase-seq, chromatin is digested by DNase I. In ATAC-seq, DNA is fragmented and tagged by Tn5 transposase simultaneously.

These methods require hundreds to millions of cells as starting material when the bench versions were initially developed. In the recent years, the sensitivities of most of the assays have been improved to single cell level, including single cell Hi-C [80], single cell ATAC-seq [30], single cell DNase-seq [81], and single-cell MNase-seq [82]. Single cell ATAC-seq was developed on a microfluidic platform. All these methods may be implemented on chip, which will dramatically improve the throughput of the assay and provide more biological insights.

In the single cell ATAC-seq [30], the assay was designed based on the C1 integrated fluidics circuit (IFC) from Fluidigm. Individual cells are captured using "butterfly" single cell trapping structures. The cells are washed and stained for viability analysis. The cell membrane is then permeabilized by surfactant NP-40, and chromatin is treated with Tn5 transposase. Open chromatin is digested and tagged by Tn5 transposase, while closed chromatin remains intact. After transposition, the Tn5-DNA complexes are dissociated by adding EDTA. By performing 8 cycles of PCR, sequencing adapters are added onto the transposed DNA. Additional PCR cycles are used to amplify libraries in 96-well plate. This method not only opened the gate for chromatin accessibility analysis using a microfluidic chip, but also demonstrated an automated, parallel library preparation protocol starting with a limited amount of DNA.

2.4 Library Construction for NGS

NGS technology has already had a profound impact on the understanding biology with sensitivity down to single cell. Unfortunately, the preparation of properly constructed sequencing libraries from DNA or RNA is still a time-consuming and tedious process. Most library preparation procedures are still manual and not suitable for high-throughput sample preparation, not to mention the high cost. Microfluidics can manipulate hundreds of reactions simultaneously, which make it possible to streamline the entire library preparation. The ability to work with a small reaction volume (µl to pl) may potentially improve the sensitivity. A typical library preparation procedure involves nucleic acid extraction, fragmentation, adapter ligation, amplification, and library quantification. Extracting high quality DNA [5, 7, 83-85] or RNA [86-90] from various species has been extensively investigated and can be found in review articles [13, 91-96]. In this review, we will not focus on nucleic acid extraction methods.

DNA fragmentation

Depending on the application, the size of DNA templates are usually at least a few thousand base pairs long, while NGS requires libraries with 200-600 bp length in order to bind to the sequencing flow cell. The DNA template needs to be fragmented before it is used for library construction. The most common methods are sonication and enzymatic fragmentation.

i. Sonication

The conventional way to fragment DNA is using sonication. The ultrasonicator employs focused bursts of ultrasonic energy to a specific focal zone where numerous

cavitation bubbles are generated. When each burst ends, these small bubbles collapse, create high velocity jets of solute, and break DNA into small fragments.

Tseng et al. described a DNA/chromatin shearing device within a microfluidic chip [97]. The acoustic field is generated by attaching a piezoelectric Langevin-type composite transducer to a microfluidic chip. The fragment size can be controlled over a range from 180 to 4000 bp by adjusting voltages and pulse duration. Our group extended the work by adding crescent shaped structures in the microfluidic chamber, which improved the generation of cavitation [27]. The sheared DNA or chromatin was used for ChIP or MeDIP analysis followed by qPCR. Such integration allowed highly sensitive assays with ~100 cells for ChIP-qPCR and 500 pg DNA for MeDIP-qPCR.

ii. Enzymatic fragmentation

Fragmentase (NEBNext, New England Biolabs) and DNase I [98] can also be used for DNA fragmentation. NEBNext dsDNA Fragmentase is currently the more popular choice. It is a mix of two enzymes and generates DNA fragments between 100-800 bp in length by adjusting incubation time. The enzyme can be inactivated by heat at 65 ℃ for 15 min. Since there is no additional equipment needed, enzyme fragmentation can be easily scaled up for high throughput library preparation. It has been employed in several applications including RNA-seq, DNA-seq, and haplotyping [99, 100]. It also showed the highest consistency among enzymatic fragmentation, sonication, and nebulization [101]. Even though applications of using Fragmentase in the microfluidic chip have not been reported, DNase has been implemented on chip [102] for HIV genotyping. The device automatically conducted RNA purification, RT-PCR, nested PCR, DNase fragmentation, and hybridization to GeneChip oligonucleotide arrays.

Ligation

In order to allow the DNA fragments to attach to the flow cell, the DNA needs to be ligated with adapters on both ends by ligase. When dealing with a limited amount of DNA, it is critical to ensure efficient ligation. Ligation between adapters (adapter dimer) reduce library quality.

Hong et al. designed the first microfluidic chip for DNA ligation, even though it is not specialized for NGS [103]. DNA, vector, and enzyme are filled in three consecutive channels. These three solutions are pushed into a mixing ring and mixed by an actuating peristaltic pump. After incubating for 15 min, the DNA ligated to vector is eluted and ready for transformation. Similarly, Lin et al. used an electrowetting-on-dielectric (EWOD) microfluidic chip for DNA ligation with vectors [104]. Reagents in their own reservoirs are separated into droplets, and the droplets are moved in the common microchannels and mixed. The solutionis manipulated by electrical potential instead of external pump or micromechanical valve. Ko et al. combined micromixer with a microchannel reaction to reduce the complexity of the EWOD system [105].

Integrated library prep

Library preparation is a time consuming and costly procedure. An automated sample preparation platform may help to reduce hands-on time and reagent cost. A key step to perform library preparation is DNA purification. The enzyme, buffer, and small molecules used in each step need to be removed to avoid interference with the following steps. Kim et al. used digital microfluidic (DMF) platform and AMPure XP magnetic beads to integrate multiple subsystem modules [106]. AMPure XP beads bind to large DNA and exclude DNA smaller than certain size, based on the buffer composition. DMF utilizes

electrode arrays to transfer and merge liquid. It is programmed to exchange the buffer of the beads and wash the beads after binding. After this demonstration, they further adapted the entire tagmentation based Nextera library preparation protocol to their platform [107]. *E. coli* genomic DNA (9 ng) is subjected to tagmentation (fragmentation and adding adapters), clean-up, PCR amplification, and size selection. The assay is finished in about 1 h with 5 min of hands-on time. Tan et al. designed the automated, multi-column chromatography (AMCC) chip to perform multiple purification on 16 independent samples [108]. ChargeSwitch beads and AMPure beads are packed into a column. Both beads are capable of capturing/releasing DNA, depending on buffer composition. Peristaltic pumps were integrated to mix the samples with buffers and force samples to flow through the column for purification. Fagmented DNA (100 ng) was end-repaired, dA tailed, ligated with adapters, and size selected on chip. The assay was finished in about 4 h with 25 min hands-on time for 16 samples.

Library quality control

Two major aspects to evaluate the quality of sequencing library are the fragment size and library concentration.

i. Library quantification

Depending on the sequencer and sequencing facility, libraries with 2-10 nM concentration are usually required for sample submission. Reliable quantification of library concentration will help to achieve the optimal amount of reads during sequencing. Spectrometer (Nanodrop), Fluorometer (Qubit) and quantitative PCR (KAPA library quantification system) have all been used for library quantification. Because only DNA that is successfully ligated on both ends can be detected by quantitative PCR.

Quantitative PCR provides better accuracy over fluorometer and spectrometer. Digital PCR, a variation of quantitative PCR, calculates the absolute number of copies of DNA [109]. The digital PCR has been conducted in either a micro-chamber [73, 110] or micro-droplet [111]. It was demonstrated by White et al. that the digital PCR shows lower variation and higher sensitivity compared to real-time PCR based assays or spectrometer based assays [109, 112]. Digital PCR for measuring DNA copy number has been intensively reviewed [113-115], so it will not be further discussed in this review.

ii. Library fragment size

To determine the effective library concentration and justify the quality of the library, it is necessary to check the library fragment size (~200-600 bp). Gel electrophoresis was the common way to determine fragment size. Because of the minimal sample consumption and fast process, microchip-based instruments (Bioanalyzer and Tapestation) are becoming more popular. Thaitrong et al. developed an automated platform for NGS quality control to improve upon these commercially available instrument [116]. The system integrates a droplet-based digital microfluidic system, capillary-based reagent delivery unit, and quantitative capillary electrophoresis module. It is capable of measuring DNA of 5-100 pg/µl and requires much less sample than Bioanalyzer.

CHAPTER 3 – Diffusion-based Microfluidic PCR for "One-pot" Analysis of Cells

3.1 Introduction

Analysis of genes from targeted cells based on polymerase chain reaction (PCR) is routinely required for illustrating the fundamental molecular biology involved in cellular events and detecting abnormal pathways involved in disease development. Microfluidic devices offer potential for genetic analysis based on tiny amounts of cell samples with high sensitivity and high degree of integration [117-124]. Genetic analysis of cells often starts with cell lysis which releases genes from cells. The released genes typically require isolation and purification before they are analyzed by PCR-based amplification [125, 126]. The integration of these steps on a microfluidic platform warrants careful consideration and arrangement. The entire process involves various chemical and biological reagents, as such some reagents may strongly interfere with the functions of others. Most notably, chemical reagents used for cell lysis (such as sodium dodecyl sulphate and Triton X-100) may inhibit PCR by reducing polymerase activity [127-129]. Furthermore, the intracellular molecules such as proteins [130], polysaccharides [131], and ions [132] (including Ca^{2+}, Fe^{3+} and EDTA) in the cell lysate may also interact with a polymerase and affect the PCR result. So far there have been three strategies to remove or alleviate the impact of cell pretreatment on PCR: 1. Institute some type of isolation step to remove lysis reagent and undesired intracellular molecules while preserving nucleic acids. This is typically done by adsorption of nucleic acids on solid surfaces (e.g. beads or matrices) while replacing the solution [118, 119, 133, 134]. The adsorption and desorption processes involved in these assays

increase the complexity of the procedure. Additional chambers and structures may also be needed on the chip to accommodate the procedure. 2. Use alternative lysis methods that interfere with PCR to a less degree than surfactants, typically in combination with dilution. Freeze-thaw[135] or heating [136, 137] has been applied to cell lysis in the context of genetic analysis. These methods are less efficient than surfactant-based lysis [138]. In a recent milestone work[136], lysis, reverse transcription, and PCR were conducted in separate connected chambers of increasing sizes so that the special lysis reagent (which could be heat deactivated) and reverse transcription product were diluted significantly before PCR was carried out. The transport of the entire reaction mixture to the next larger chamber in each step ensured that the nucleic acids did not get lost. The limitation of such design is that the complexity of the device increases with the number of steps and reagents required by the analysis. 3. Use Direct PCR kit based on Phusion polymerase that is tolerant to surfactant-based lysis reagents [139]. However, Phusion polymerase can be inhibited by SYBR green and other dyes [140]and lacks 5' to 3' exonuclease activity [141]. Thus commonly-used fluorescence-based quantification is impossible with the Phusion polymerase system. To summarize, microfluidic strategies that permit simple operation and device design and are compatible with *Taq* polymerase are still in high demand. Such strategy will be particularly beneficial for precise and demanding operations such as genetic analysis based on single cells [142-144].

In this work, we demonstrate a simple scheme for conducting microfluidic PCR starting from cells, taking advantage of the difference in the diffusivity between DNA and various reagent/intracellular molecules [145]. We use the same reaction chamber for both cell lysis and PCR and this is analogous to "one-pot synthesis" (i.e. using one reactor for

successive reactions without separation and purification). The lysis buffer and the PCR mix were introduced into the chamber by concentration-gradient-driven diffusion. During such diffusion, the new solution replaces the solution and molecules from the previous step without removing the slow-diffusing large DNA molecules. We show that this scheme allows highly efficient genetic analysis of a low number of cells (including single cells). The single chamber ("one-pot") design drastically minimizes the complexity of the microfluidic device. We envision that this may be a general approach for on-chip multi-step assays on sizable DNAs.

3.2 Results and Discussion

Design for microfluidic diffusion-based PCR.

Our microfluidic device had a simple structure that included a reaction chamber connected with two loading chambers on both sides (**Fig. 3.1a**). The connections between the reaction chamber and the two loading chambers (i.e. a number of channels) could be cut off by closing two-layer valves [146, 147]. A hydration line structure was also added in the control layer so that pressure and water inside the hydration line minimized the
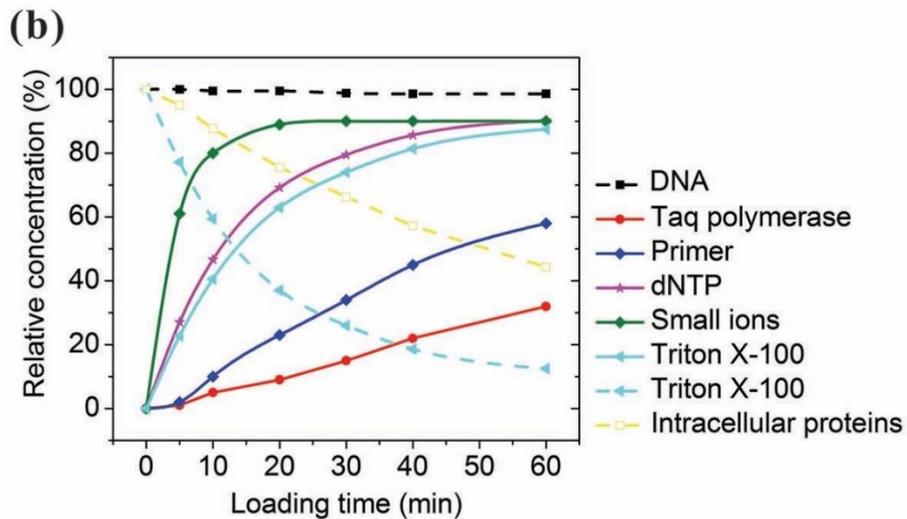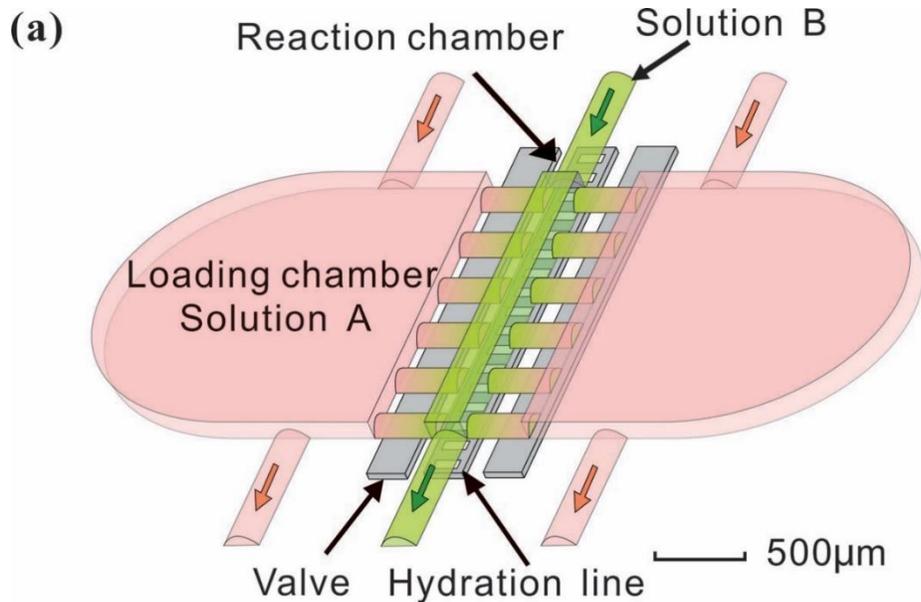
Fig. 3.1 Principle of solution replacement based on diffusion. (a) The design of the microfluidic device. The device consists of two PDMS layers. Control and hydration are implemented in the same control layer by actuating the valves and supplying water under pressure (~25 Psi). The fluidic layer has a small reaction chamber connected with two large loading chambers. When the valves are open, molecules in the solution A may enter the reaction chamber by diffusion, effectively replacing solution B. (b) The entry (solid lines) and release (broken lines) of various molecules into/out of the reaction chamber, modeled by COMSOL Multiphysics. The initial concentration in the loading chamber (in the cases of entry into the reaction chamber) or that in the reaction chamber (in the case of release out of the reaction chamber) was used as the reference (i.e. having a value of 100%).

evaporation during PCR [118, 148]. With the volume of the loading chambers substantially larger than that of the reaction chamber (by a factor of 16), this simple microfluidic platform allowed the replacement of small molecules (either reagents or intracellular molecules) inside the reaction chamber by diffusion from/into the loading chambers under concentration gradient, while preserving large DNAs inside throughout such process. The depth of the fluidic chambers is ~60 μm. It allows single cells to pass through without noticeable blocking issue. The depth also guarantees the diffusion process to reach equilibrium rapidly. Depth higher than 60 μm may help alleviate water evaporation during heating, but it would take longer for diffusion. The diffusion process is governed by Fick's first Law:

$$J = -D\frac{\partial \varphi}{\partial x} \qquad \text{Eqn 1}$$

Where J is the diffusion flux, D is the diffusivity or diffusion coefficient, φ is the concentration, and $\frac{\partial \varphi}{\partial x}$ is the concentration gradient. There is substantial difference among the diffusivity of involved species: at 25 ℃, the diffusivity of genomic DNA fragments is estimated to be $2.0 \times 10^{-13} m^2/s$ (by assuming ~50 kb as the average size and calculating using the empirical equation[149]), compared to 4.7×10^{-11} m^2/s for *Taq* polymerase [150], 8.0×10^{-11} m^2/s for intracellular proteins [151] (considering ~53kDa as the average size), 1.0×10^{-10} m^2/s for primers (~20bp ssDNA) [152], 3.7×10^{-10} m^2/s for dNTP [153], 3.0×10^{-10} m^2/s for Triton X-100 [154] and 1.0×10^{-9} m^2/s for small ions [155], such as Mg^{2+}, K$^+$ and Cl$^-$. **Fig. 3.1b** shows a COMSOL model of our microfluidic design. Only 1.4% of genomic DNA fragments diffuse out of the reaction chamber in 1 h. In comparison, 87.5% of the lysis reagent Triton X-100 and 44.3% of the intracellular proteins, diffuses out of the reaction chamber (into the loading chambers) in the same period. 1 h loading of new reagents causes the concentrations of *Taq* polymerase and primers in the reaction chamber to reach 32% and 58%, respectively, of the concentrations in the loading chambers, compared to 90% for small ions. Thus the concentrations in the loading chambers may need to be high in order to supply the reaction chamber with desired concentrations given a short loading period. It is worth emphasizing that the modeling was conducted by assuming the size of DNA to be ~50 kb which was the typical size of genomic DNA after lysis with shearing, as opposed to the size of a complete genomic DNA. Larger DNA sizes will further facilitate our scheme.

On-chip amplification of purified DNA.

As an initial test, we used our platform for PCR of purified human genomic DNA produced from a lymphoblastoid cell line (GM 12878). In the experiment, we first loaded

genomic DNA in water into the reaction chamber (about 500 copies of DNA molecules in the 24 nl reaction chamber). We then filled the loading chambers with PCR mix having designed concentrations (250 U/ml *Taq* polymerase, 1.2 μM for each primer, 0.4 mM for each of dNTPs, 6 mM $MgCl_2$ and 100mM KCl, 0.8% PEG 8000, 0.08% Tween-20) and allowed various periods for the loading times (0-30 min). It is worth noting that the concentrations of *Taq* polymerase and primers in the loading chamber were significantly higher than the desired concentrations (50 U/ml and 0.4 μM, respectively) in order to speed up the diffusion-based material transfer. During loading, the PCR mix (with primers targeting GAPDH gene) slowly replaced the original solution in the reaction chamber by diffusion. The reaction chamber was then sealed off by the valves and the microfluidic chip was placed on a flat-plate thermal cycler for PCR amplification. After the PCR amplification, the PCR product was quantified using qPCR to obtain the copy number of the amplified genes. As shown in **Fig. 3.2**, we observed increasing amplification with longer loading times (the grey bars). To verify that the PCR result was dictated by the diffusion process, we used COMSOL to model the molecular transport during loading and generated values for *Taq* polymerase (0, 0.2, 2, 13 and 38 U/ml) and primers (0, 3, 20, 120, 400 nM) in the reaction chamber under various loading times (0, 1, 3,10, 30 min). We then conducted on-chip PCR in the reaction chamber containing PCR mixes with the above *Taq* polymerase/primer concentrations (striped bars). We found that the amount of PCR product was very similar in these two cases. This confirms that the increase in the amplification under longer loading time was due to higher reagent concentrations in the reaction chamber after longer diffusion and the diffusion of *Taq* polymerase and primers was the limiting step.
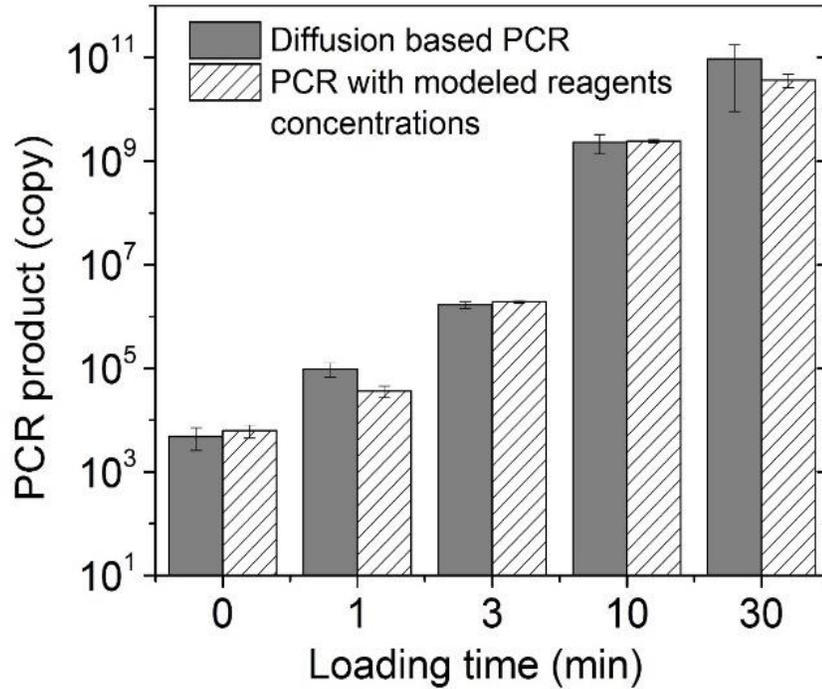
Fig. 3.2 Diffusion-based PCR in a 24nl microfluidic chamber starting with genomic DNA purified from GM12878 cells. The detection targeted GAPDH gene and the amplification run for 30 cycles. The copy number of the PCR product was quantified using qPCR. The results with various loading times (grey bars, 0, 1, 3, 10 and 30 min) are compared with those of on-chip PCR with various Taq polymerase concentrations (0, 0.2, 2, 13 and 38 U/ml) and primers concentrations (0, 3, 20, 120, 400 nM) (striped bars, these values were generated by COMSOL modelling of the diffusion process for the corresponding loading times) to confirm the impact of the diffusion on PCR results.
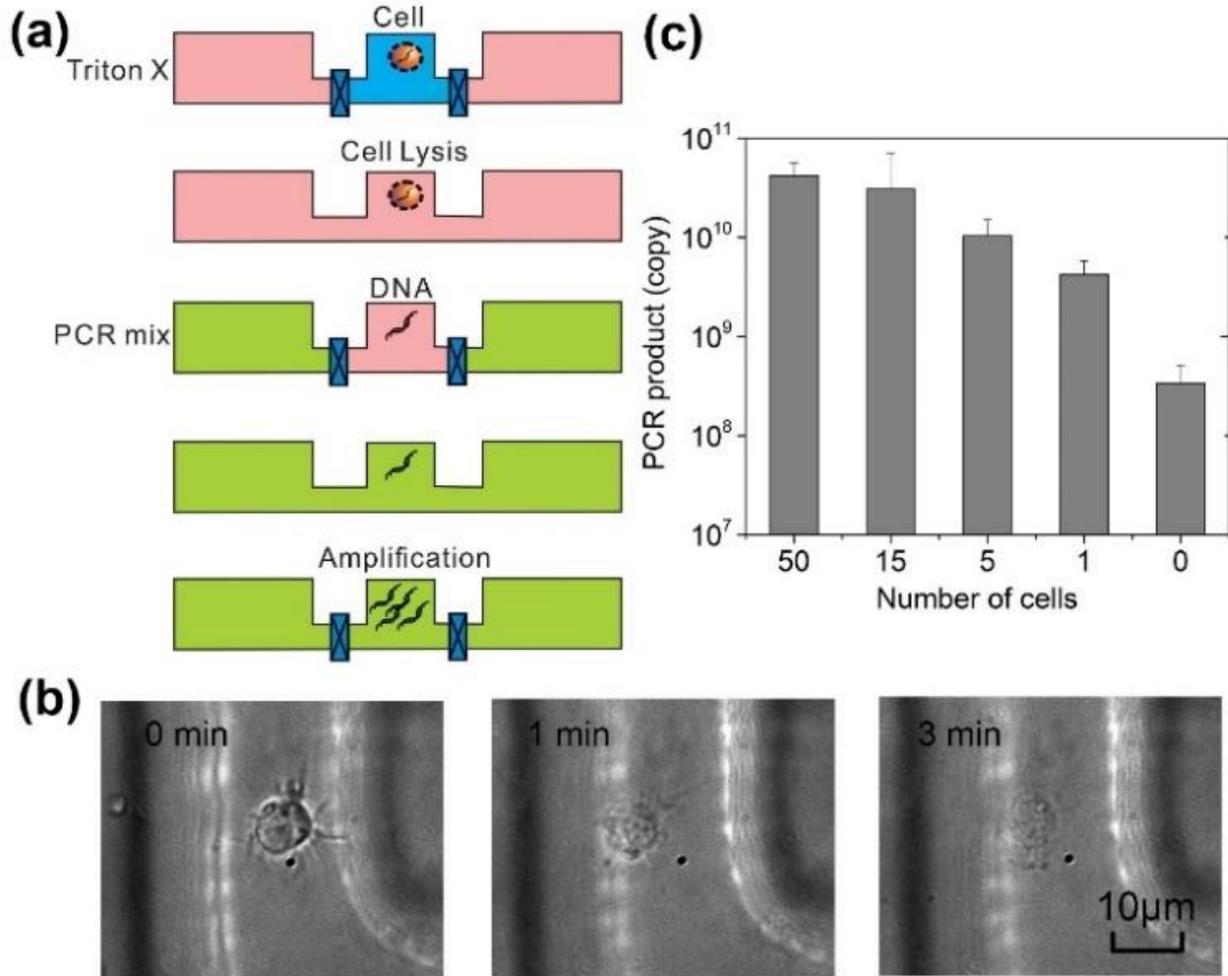
Fig. 3.3 Combined lysis and PCR starting with GM12878 cells in a 3nl reaction chamber. (a) A schematic on the procedure. Triton X lysis buffer replaces the original cell buffer by diffusion. Then the triton X lysis reagents are replaced by PCR mix by diffusion. Finally, PCR amplification occurs for 45 cycles. Triton X lysis buffer and PCR mix had loading times of 10 and 30 min, respectively. (b) The cell lysis observed under differential interference contrast (DIC) microscope. Cells are completely lysed after the loading of Triton X lysis buffer for 3 min. (c) The combined lysis and PCR procedure on cells of various numbers (0-50). The PCR signal resulted from the control sample (0 cells) is from primer dimerization as shown by the melting curves (**Fig. 3.5**).
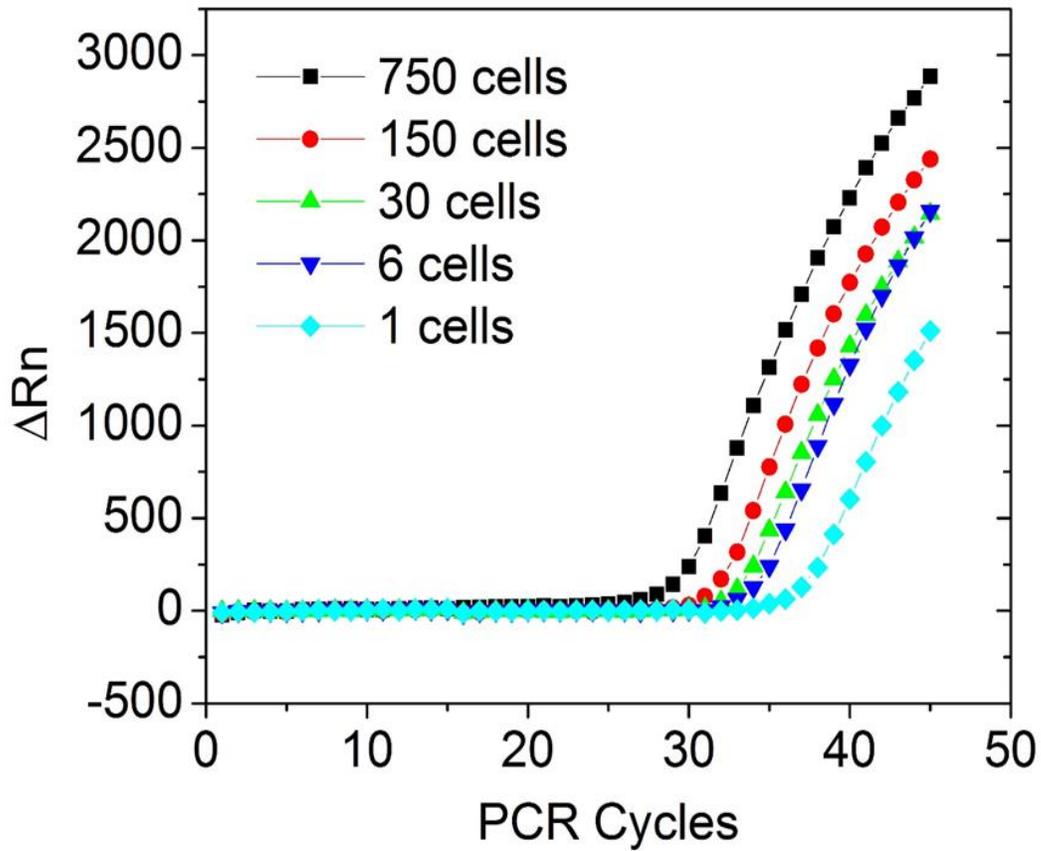
Fig. 3.4 Real-time PCR amplification of purified DNA from various numbers of GM12878 cells (purified by QIAamp DNA Blood Mini Kit). There is noticeable difference in the signal when the number of the cells varies.
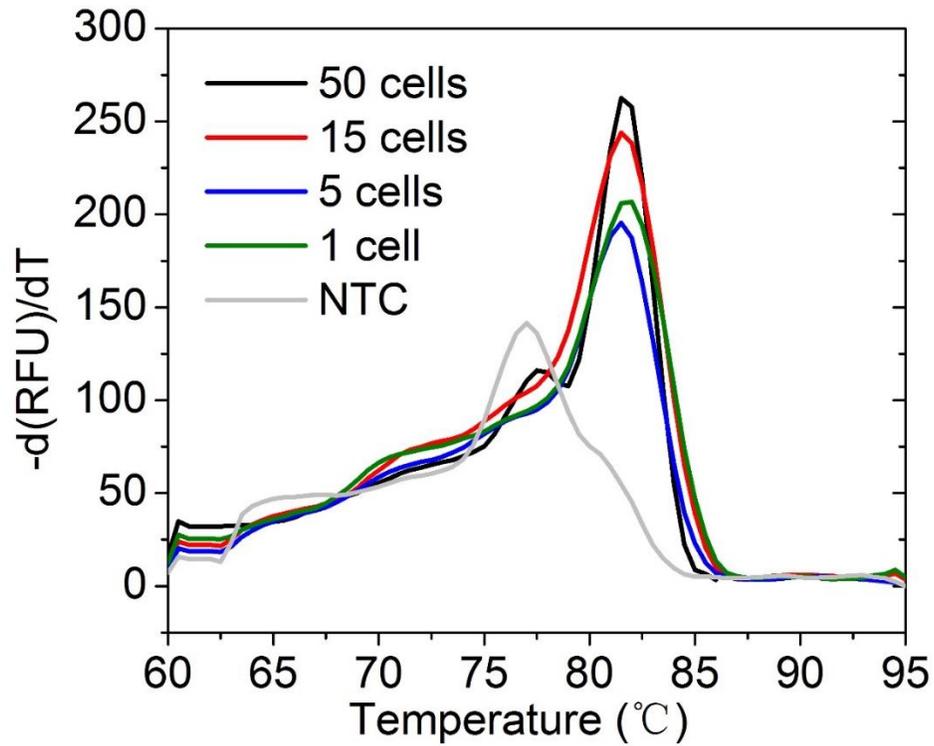
Fig. 3.5 The melting curves of the on-chip samples in Fig. 3.3. Amplification products were eluted from the chip and measured by real time PCR. NTC is the no template control (negative control).

<u>*On-chip amplification starting from cells.*</u>

Based on this understanding, we designed the multi-step PCR process, starting from cells (**Fig. 3.3a**). Compared to the experiments in **Fig. 3.2**, we reduced the volume of the reaction chamber to 3 nl and used it to process 1-100 cells and applied 45 thermal cycles for the amplification. GM12878 cells were first loaded into the reaction chamber. Triton X was flowed into the loading chambers (with the reaction chamber closed) and then diffused into the reaction chamber for lysing the cells with a loading time of 10 min (shown in **Fig. 3.3b**). The reaction chamber was then closed for 5 min for lysis to complete
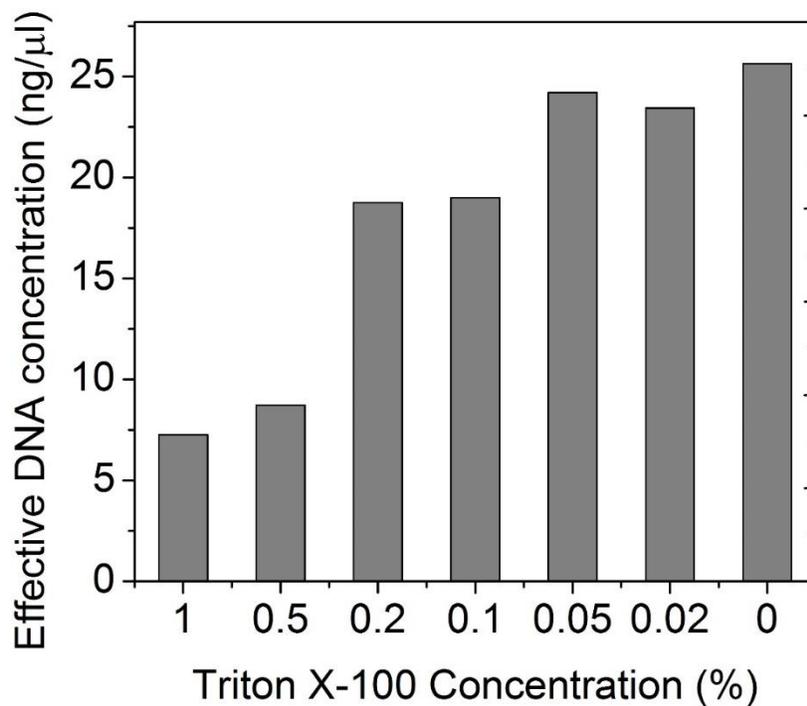
Fig. 3.6 The inhibition of off-chip PCR by Triton X-100. 25 ng/µl DNA template with various Triton X-100 concentrations in each test was measured by real time PCR. The measured amount of DNA (i.e. effective DNA concentration) is lower than the actual amount (25 ng/µl) due to inhibition of Triton X-100.

while the PCR mix (250 U/ml *Taq* polymerase, 1.2 µM for each primer, 0.4 mM each of dNTPs, 6 mM MgCl$_2$ and 100 mM KCl, 0.8% PEG 8000, 0.08% Tween-20) was flowed into the loading chambers. The PCR reagents were then diffused into the reaction chamber with a loading time of 30 min, while the lysis reagent triton X and intracellular molecules diffused out. We run the thermal cycles (45 cycles) while the reaction chamber was closed. The PCR product was then flushed out of the reaction chamber for quantification by qPCR off-chip. **Fig. 3.3c** shows the results of the diffusion-based PCR for detecting samples ranging from single cells to 50 cells. Real-time PCR used in our experiment has the resolution to differentiate various numbers of cells in the range of 1-1000 (as shown in **Fig. 3.4**). **Fig. 3.3c** shows that various numbers of cells (1, 5, 15 and 50) were easily differentiated based on the amount of PCR product after amplification of 45 cycles. It is worth noting that the signal from the blank sample (0 cell) was due to the dimerization of primers as suggested by the melting curves (**Fig. 3.5**). The parameters of the above process were carefully designed to maximize the efficiency for the on-chip PCR. Our modeling shows that the residual Triton X in the reaction chamber during PCR was less than 0.1% which has minimal effect on PCR performance based our test results (**Fig. 3.6**). As a further confirmation, the diffusion-based PCR starting with cells yielded similar results compared to those starting with purified DNA from the same number of cells (**Fig. 3.7**), indicating that there was no inhibition from the lysis reagent or intracellular molecules.

3.3 Materials and Methods

Microfluidic chip fabrication
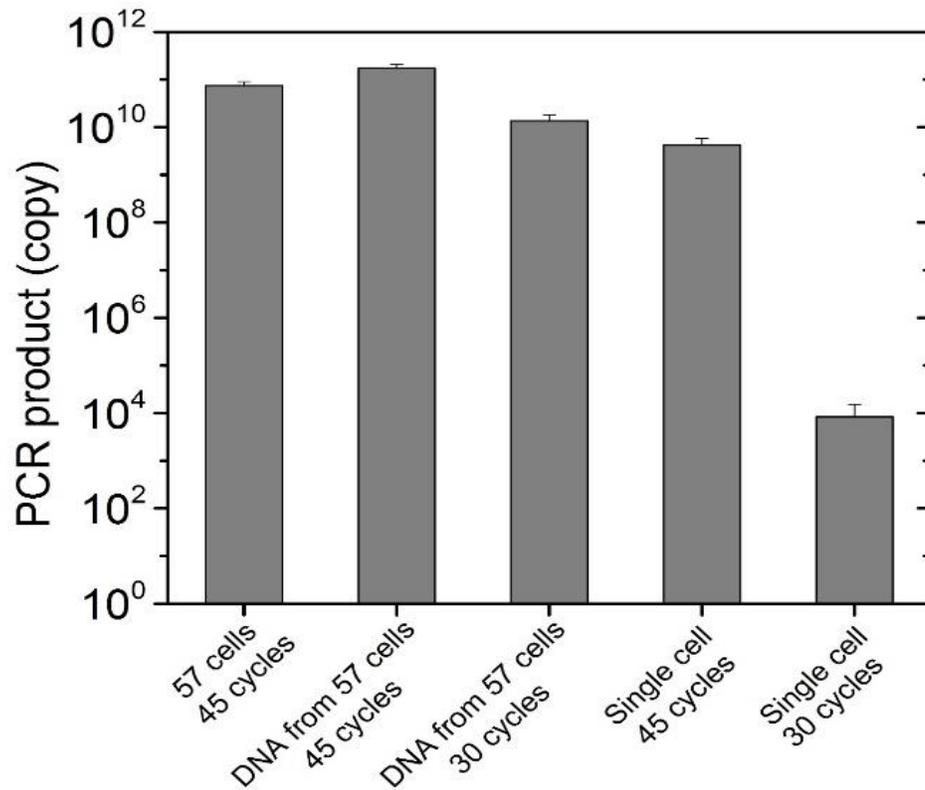
Fig. 3.7 The comparison among on-chip PCR tests (with 3nl reaction chamber) of various conditions. Some of these tests were conducted with combined lysis and PCR process (Loading times for Triton-X and PCR mix were 10 and 30 min, respectively), starting from GM12878 cells (e.g. "57 cells"). The others were done with purified DNA from a certain number of GM12878 cells (e.g. "DNA from 57 cells"), with the purification conducted using QIAamp DNA Blood Mini Kit. Samples were amplified for either 30 or 45 cycles.

Two-layer PDMS chip was fabricated by multilayer soft lithography [147] (**Fig. 3.1a**). Photomasks were designed by Freehand MX (Macromedia, San Francisco, CA) and printed on 4000 dpi films. The fluidic master (photoresist on a silicon wafer) was fabricated in SU-8 2025 (Microchem, Newton, MA) and AZ 9260

(Clariant, Charlotte, NC) with the thickness being 60 μm and 13 μm, respectively. The master was heated to 130℃ to form rounded cross-sectional profile for the features in AZ 9260. The control layer master was fabricated in SU-2025 with 24μm thickness. The control layer PDMS was made by spinning PDMS (RTV615A: RTV615B=20:1, R. S. Hughes, Sunnyvale, CA) at 500 rpm for 10s and then at 1700 rpm 30s, which resulted a thickness of 67 μm. The fluidic layer PDMS had the composition of RTV615A: RTV615B =5:1 and a thickness of ~0.4 cm. Both layers of PDMS were cured at 80℃ for 15 min. The two layers were then aligned, brought into contact and baked for 1h at 80℃. The two-layer PDMS structure was then peeled off from the master and access holes were punched. Finally, the PDMS structure was bonded to clean cover glass (~150 μm thick, VWR, Radnor, PA) after plasma oxidation of the surfaces (Harrick Plasma, Ithaca, NY) and baked at 80℃ overnight to strengthen the device.

Device setup and operation

The solutions were delivered into the device by a syringe pump. The actuation of the two-layer valves and hydration line was controlled via solenoid valves (ASCO Scientific, Florham Park, NJ) using a DAQ card (NI SCB-68) and LabVIEW (National Instruments, Austin, TX) programs [26, 83]. The control layer and the hydration line were filled with water before experiments.

Human genomic DNA purified from a lymphoblastoid cell line (GM 12878) using a QIAamp DNA Blood Mini Kit (Qiagen, Valencia, CA) were used in on-chip PCR testing. During the PCR, the loading chambers were filled with PCR mix of a designed composition (1x Fast Plus EvaGreen qPCR Master Mix from Biotium,

0.8% PEG 8000, 0.08% Tween-20, 0.2U/µl Taq polymerase, 1.2mM GAPDH primers for GM 12878 cells). The microvalves were open for 10 to 30 min to replace the original solution in the reaction chamber with desired reagent concentrations. The microfluidic device was then placed on a flat-plate thermal cycler (Techne, Bibby Scientific, Burlington, NJ) for on-chip PCR. A metal frame was used to clamp down on the microfluidic chip to ensure good contact with the flat plate. Paraffin oil (Sigma-Aldrich, St. Louis, MO) was applied in between the glass surface of the chip and the flat plate to improve heat transfer. Thermocycling conditions were 95℃ for 3min followed by 30 cycles or 45 cycles of (95℃ for 30 s, 60℃ for 30 s and 72℃ for 30 s).

When cell-lysis was included as an on-chip step, GM 12878 cells (suspended in a PBS buffer at a concentration of 10^6 to 10^7 cells/ml) flowed into the device to leave a known number of cells in the reaction chamber. A Triton X lysis buffer (1% Triton X-100, 50mM Tris-HCl, 150mM NaCl, pH 7.5) was flowed into the loading chambers. Opening the microvalves for 10 min permitted complete cell lysis. The cell lysis step was then followed by diffusion-based PCR described above.

Quantitative PCR assay on the on-chip PCR product

QPCR assays were performed in 25µl aliquots in a Real-Time PCR Detection System (CFX Connect, BIO-RAD, Hercules, CA). The product from on-chip PCR was eluted by 10 µl H_2O and then serial- diluted to the working range of the qPCR. Thermocycling conditions were 95℃ for 3min followed by 30 cycles of (95℃ for 30 s, 60℃ for 30 s and 72℃ for 30 s). The following primers were used to detect GAPDH gene in GM 12878:

GAPDH forward: 5'- CCCCACACACATGCACTTACC -3',

GAPDH reverse: 5'- CCTACTCCCAGGGCTTTGATT -3'.

Mathematical modelling

COMSOL Multiphysics 4.3 was used to model the diffusion process in the microfluidic device (i.e. molecular exchange between the loading chambers and the reaction chamber). The physics model, namely transport of diluted species was used. A 3D time-dependent study was built to analyse the concentration variations of each species within 1h inside the microscale structure. No flux at all the boundaries was applied. The average concentration of each species in the reaction chamber was evaluated. We modelled the diffusion processes at 25 ℃ using the diffusivity values included in the main text. To model the entry of species (Triton-X 100, dNTP, Taq polymerase, primers, and small ions) from the loading chamber into the reaction chamber, the starting concentration in the reaction chamber was 0 and then we examined the increase in the species concentration in the reaction chamber over time in terms of its percentage of the original concentration in the loading chamber. To model the exit of the species (DNA, proteins, and Triton-X 100) out of the reaction chamber, the starting concentration in the loading chamber was 0 and then we examined the decrease in the species concentration in the reaction chamber over time in terms of its percentage of the original concentration in the reaction chamber.

3.4 Conclusion

Genetic analysis starting with cell samples often requires multi-step processing including cell lysis, DNA isolation/purification, and polymerase chain reaction (PCR) based assays. When conducted on a microfluidic platform, the compatibility among

various steps often demands complicated procedure and complex device structure. Here we present a microfluidic device that permits "one-pot" strategy for multi-step PCR analysis starting from cells. Taking advantage of the diffusivity difference, we replace the smaller molecules in the reaction chamber by diffusion while retaining DNA molecules inside. This simple scheme effectively removes reagents from the previous step to avoid interference and thus permits multi-step processing in the same reaction chamber. Our approach shows high efficiency for PCR and potential for a wide range of genetic analysis including assays based on single cells.

The use of diffusion for changing small-molecule reagents in a microfluidic reactor is universally applicable to assays involving large DNA molecules isolated from cells. Our approach efficiently eliminates DNA isolation and purification steps and drastically simplifies the design of the microfluidic device. These advantages will be critical for assays starting from a low number of cells or processes involving large-scale parallel operations (e.g. analysis of single cell arrays). The major drawback of the approach is the added time required for diffusion. This additional assay time can be shortened by elevation of the temperature and increase in the concentration in the loading chamber.

CHAPTER 4 – Cell Type-specific Mouse Brain Methylomes Profiled Using an Ultralow-input Microfluidic Device

4.1 Introduction

DNA methylation patterns in the genome (i.e. DNA methylomes) critically affect gene activities as a part of overall epigenetic regulatory program during normal development and disease processes. CpG dinucleotides are important targets for methylation (mCG). Hypermethylation on CpG island promoters is a common mechanism for gene silencing and an epigenetic feature for many types of human cancers [156]. Both mCG [157, 158] and non-CG methylation (mCH, where H = A, C, or T) [21, 159] are critically involved in human neuronal genome and their dynamics play critical roles in mammalian brain development. There have been a number of genome-wide technologies for profiling DNA methylomes. Whole-genome bisulfite sequencing (WGBS) is generally considered the gold standard for DNA methylation analyses [49, 160]. Bisulfite treatment converts cytosine residues to uracils, with 5-methylcytosine residues unaffected. Combined with high-throughput sequencing, the approach generates methylomic profiles with single-nucleotide resolution. Despite of the high resolution, the cost associated with deep sequencing involved in WGBS (>500 million reads per sample to cover human genome) can be prohibitive. In comparison, enrichment-based technologies dramatically reduce the required sequencing depth by enriching methylated DNA fragments based on affinity purifications, at a price of low resolution (100-300 bp) [63, 161]. As another cost-effective alternative to WGBS, reduced representative bisulfite sequencing (RRBS) utilizes methylation-insensitive restriction enzyme (MspI) digestion and size selection to enrich a

subset of the genome (mostly CpG islands and promoter regions) for analysis while preserving single-nucleotide resolution [47, 162].

Prevailing WGBS and RRBS protocols require substantial amounts of genomic DNA. WGBS requires 10 ng to 5 µg input DNA [49, 160, 163]. RRBS typically requires >300 ng DNA to detect 1.0-2.2 million unique CpGs (≥1× coverage) and merely 0.3-1.0 million CpGs ≥10× coverage in human genome. Recent efforts on low-input RRBS lowered the starting DNA amount to 30 ng with compromised CpG recovery (0.9 million unique CpGs) [162]. Single-cell assays with RRBS [50] and WGBS [51, 164] were demonstrated recently. In spite of discovering variations among individual cells, single cell RRBS or WGBS data covered only a fraction of the genome (up to 48%) [50, 51, 164]. Thus low-input methods with high methylome coverage are desired for profiling scarce samples including those derived from animals and patients.

Here we demonstrate a microfluidic technology, referred to as MIcrofluidic Diffusion-based RRBS (MID-RRBS), for low-input (down to 0.3-0.5 ng DNA) assay with high bisulfite conversion efficiency and high coverage of CpGs. We used a diffusion-based reagent swapping approach for the multi-step treatment of DNA on the microfluidic platform. Our microfluidic bisulfite conversion protocol preserved substantially more amplifiable DNA than conventional bisulfite treatment [47, 165-167] while achieving high conversion rate. We mapped methylomes of a cell line (GM 12878), neurons and glia isolated from mouse cerebellums. Our protocol yielded data that were heavy in CpGs of high coverage (2.0-2.7 million CpGs with ≥1×, 1.3-1.8 million CpGs with ≥10× coverage with 0.3 to 10 ng starting DNA, respectively). MID-RRBS technology allowed differentiation of methylomic landscapes of neurons and glia from mouse cerebellum,

generating insights into cell type-specific features and their relationships to gene expressions.

4.2 Results and Discussion

Device and protocol for MID-RRBS

Our microfluidic device consisted of a reaction chamber (~240 nl in volume) connected with two loading chambers (~480 nl each) on both sides (**Fig. 4. 1a and Fig. 2a)**. The reaction chamber and the loading chambers could be disconnected by actuating a pair of pneumatic valves in between them. We conducted a bisulfite conversion process using a diffusion-based reagent swapping method. This approach takes advantage of the difference in the diffusivities of DNA and small molecules. DNA of substantial sizes have much smaller diffusivity than small-molecule reagents involved in DNA treatment. When material exchange occurs only by diffusion, DNA loss from the small reaction chamber during a short period is minimal, whereas a small-molecule reagent in the same chamber gets released (into the loading chambers) and replaced by another small-molecule reagent needed for the next step. Although we had applied similar principle to process large DNA fragments (~50 kb) in previous work [9], treating smaller DNA fragments (several hundred bp, which is the size range relevant to next-generation sequencing) remained challenging. In our process (**Fig. 4. 1b**), a genomic DNA sample (containing fragments ligated to adapters with the sizes of 160-360 bp) was loaded into the reaction chamber and a bisulfite conversion buffer was loaded in the loading chambers, while the valves between the chambers were closed. The dosing of bisulfite mix into the reaction chamber then occurred by opening the valves and allowing diffusion under concentration gradients

(in 3 periods of 10 min each, with the loading chambers replenished with fresh bisulfite mix after each period). The diffusivity of DNA fragments (e.g. 2×10^{-11} m^2/s for ~250 bp DNA) differs with those of small-molecule reagents (~10^{-9} m^2/s) roughly by 2 orders of magnitude. Thus the majority of the DNA molecules (93% estimated by COMSOL modeling in **Fig. 4. 1c**) remained in the reaction chamber, while the small-molecule bisulfite mix filled up the reaction chamber to reach 98% and 97% of the loading concentrations for sodium bisulfite and quinol, respectively (**Fig. 4. 1c**). After sulphonation of 1 h at 55 ℃ with all valves closed, using the same diffusion-based swapping scheme, NaOH was loaded into the reaction chamber to replace bisulfite mix for desulphonation. The valves were then closed for 20 min to allow desulphonation to take place in the reaction chamber. DNA was finally eluted out of the device for ethanol precipitation and downstream sequencing. It is worth noting that our process eliminated column-based purification between bisulfite conversion and desulphonation (which may cause loss of 80-90% of DNA [167]). Small-molecule reagents in the reaction chamber were effectively removed by the diffusion step (e.g. the concentration of sodium bisulfite in the reaction chamber fell to 2% of its original one after three 10-min periods of diffusion, **Fig. 4. 1c**).
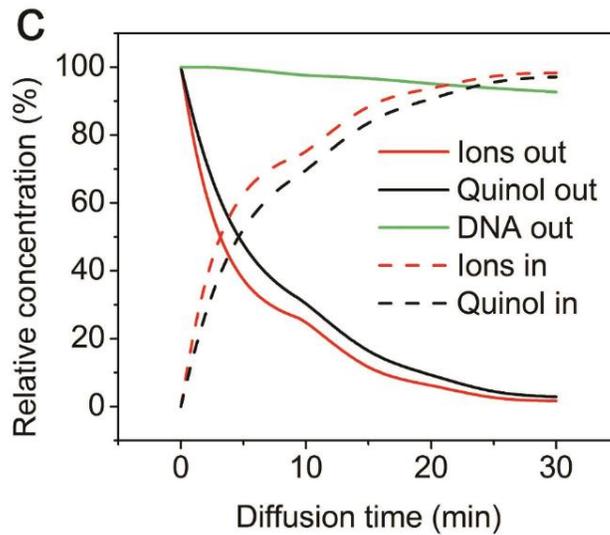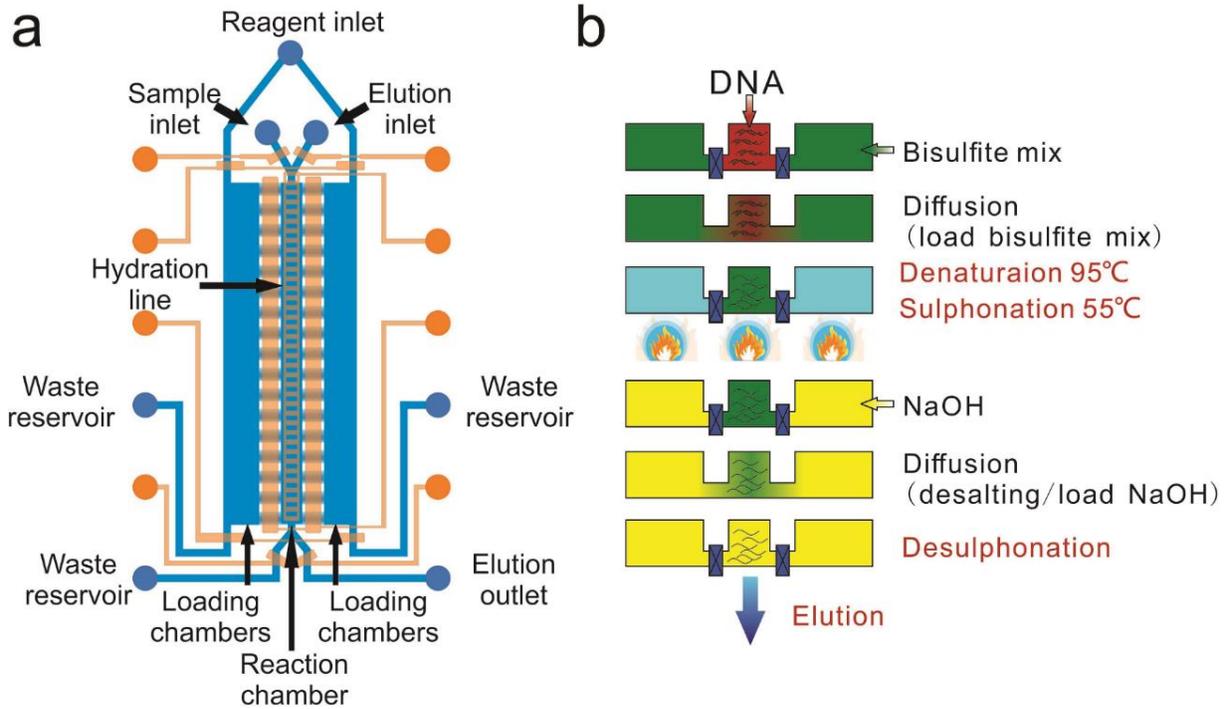
Fig. 4.1 MID-RRBS device and protocol. (a) A schematic illustration of the two-layered microfluidic device (with the fluidic layer labeled blue and the control layer labeled orange). The device consisted of one reaction chamber (at the center) and two connected loading chambers. A hydration line (within the control layer) containing water was pressurized during heating to minimize water loss from the reaction chamber. (b) Steps involved in MID-RRBS (detailed in Materials and Methods). The schematics were drawn with a cross-sectional view of the three chambers. (c) COMSOL modelling of release (solid lines) and loading (broken lines) of various molecular species from/into the reaction chamber. The release and loading of the molecules were conducted in three diffusion periods of 10 min each with the solution in the loading chambers refreshed after each period. The initial concentration in the loading chambers was considered as 100% in the case of loading whereas the one in the reaction chamber as 100% for release.

To benchmark the performance of MID-RRBS, we used our technology to profile the methylome of GM12878 cells (a lymphoblastoid cell line). We varied the amount of starting genomic DNA in the range of 0.3 to 10 ng. After MspI digestion and adapter ligation, DNA fragments in the range of 160-360 bp (with insert sizes of 40-240 bp) were selected by gel-free size selection (**Fig. 4. 3**) [168]. MID-RRBS was conducted and sequencing data were analyzed. We achieved averagely 99.8 ± 0.1% conversion rates using our protocol (**Table 4. 1**). The numbers of unique CpGs ranged from 2.0 million with 0.3 ng DNA to 2.6 million with 10 ng DNA whereas CpGs with ≥10× coverage ranged from 1.3 million with 0.3 ng to 1.8 million with 10 ng DNA (**Fig. 4. 4a**). The samples of 10,

3, 1, and 0.3 ng had Pearson coefficients of 0.95, 0.98, 0.96, and 0.84 between replicates, respectively (**Fig. 4. 4b**). We also examined the breath of coverage of various genomic regions by data obtained with various sample sizes (**Fig. 4. 4c**). Although there was gradual decline in the breath of coverage for all genomic regions with decreased DNA amount, low-input samples (1 ng or less) preserved most of the coverage achieved by 10 ng samples. For example, minor decreases from 66.6% to 63.7% with ≥1× CpG measurements and 64.5% to 61.1% with ≥5× CpG measurements were observed for core promoters when the sample size changed from 10 ng to 0.3 ng. Compared to ENCODE data (obtained using 1 µg DNA), our data contained a much larger percentage of these genomic regions with ≥100× CpG measurements. For example, the percentage of core promoters covered by ≥100× CpG measurements was 53.8% at 10 ng, 48.4% at 1 ng and 47.4% at 0.3 ng, compared to 39.7% with ENCODE data. Our aligned reads for various samples (0.3-10 ng) exhibited higher complexity than that of ENCODE data (**Fig. 4. 5**). We also compared MID-RRBS data with ENCODE data and other published data on mammalian cell lines other than GM12878 [50, 168, 169], in terms of the number of CpGs at various coverages (**Fig. 4. 4d**) and saturation analyses (**Fig. 4. 4e**). All of our data (0.3-10 ng) performed better than ENCODE (obtained with 1 µg) data in the entire spectrum. Our data yielded a much higher number of CpGs with high coverage (>30×) than mRRBS data (obtained with 100 ng using gel-free method) (**Fig. 4. 4d**). Saturation analysis revealed that our datasets taken with ≥1 ng DNA yielded higher or comparable CpG detection efficiency at all sequencing depths compared to competing works with ≥100 ng input DNA [168, 169] (**Fig. 4. 4e**). Our 0.3 ng samples (equivalent to DNA from 50 cells)

yielded far superior CpG recovery than single cell RRBS [50] and were similar to that of mRRBS, in terms of CpG detection efficiency at a given sequencing depth (**Fig. 4. 4e**).
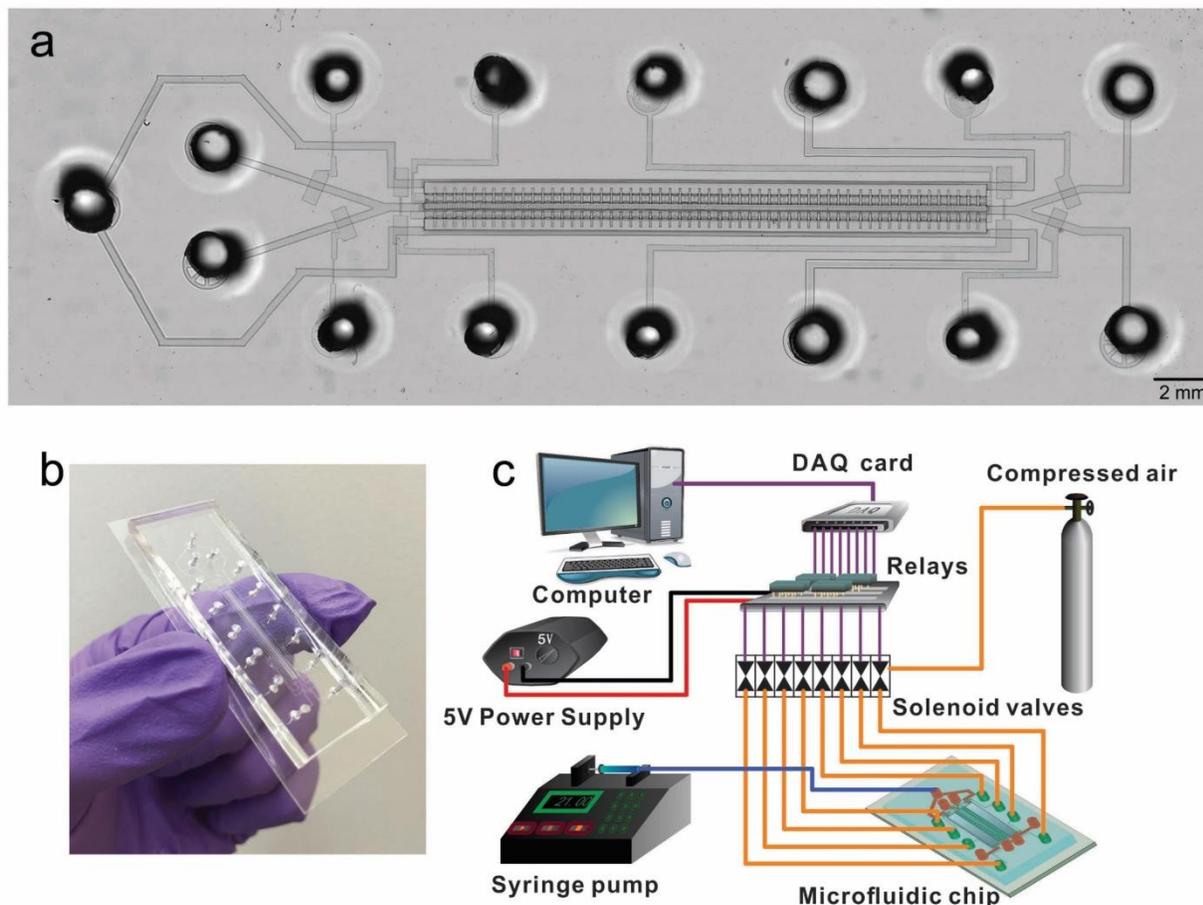
Fig. 4.2 Design and setup of the microfluidic device. (a) A microscopic image of the device (stitched from multiple images). The dimensions of the reaction chamber and each loading chamber were 20 mm × 0.2 mm × 0.06 mm and 20 mm × 0.4 mm × 0.06 mm, respectively. (b) A photograph of the PDMS/glass microfluidic device. (c) A schematic of the ancillary system for control and liquid delivery and its connection with the microfluidic device.
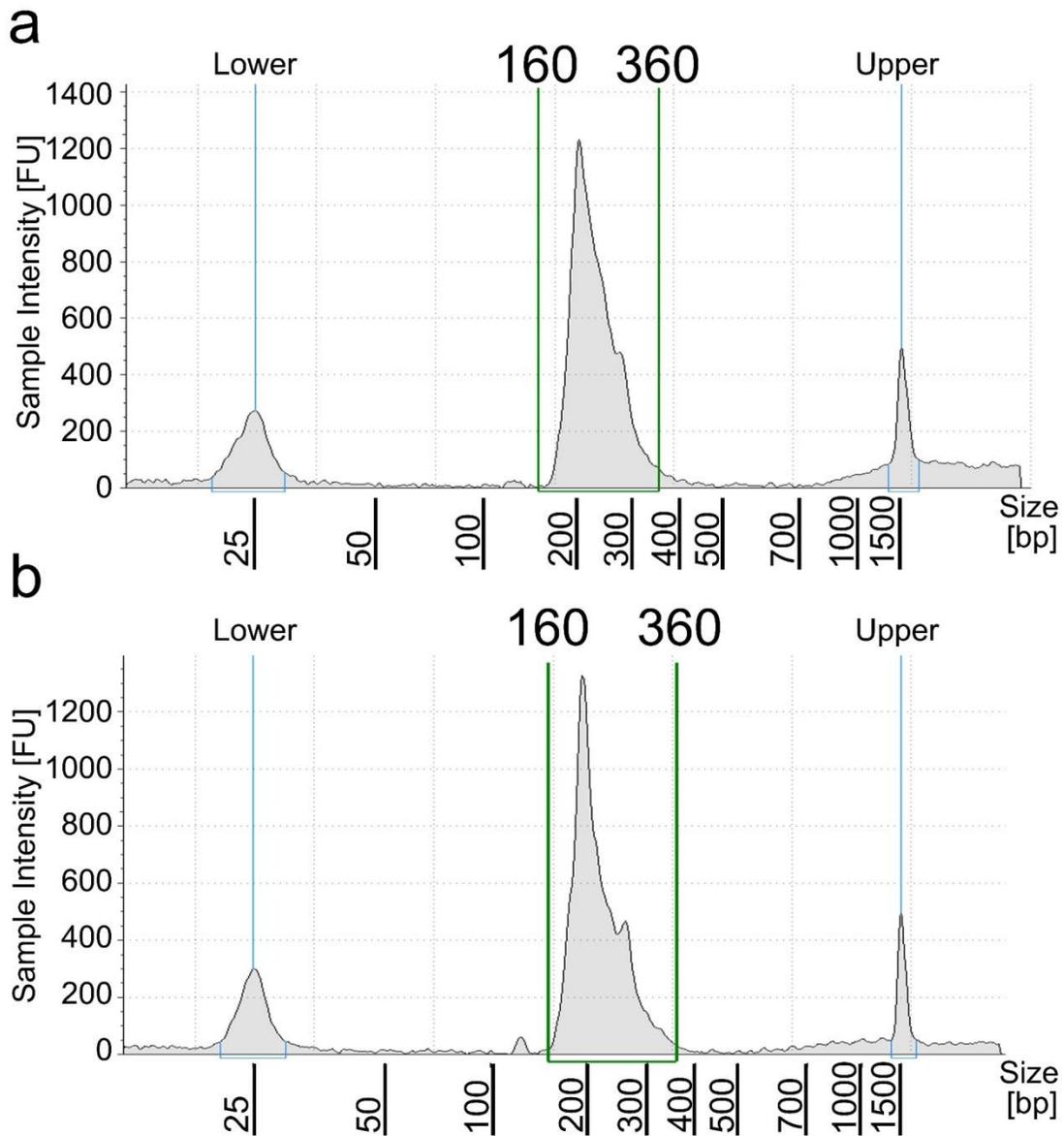
Fig. 4.3 Fragment size distributions of MID-RRBS libraries produced by gel size selection and gel-free size selection (DNA from GM 12878). (a) The fragment size distribution of a library prepared using gel size selection. (b) The fragment size distribution of a library prepared using gel-free size selection.
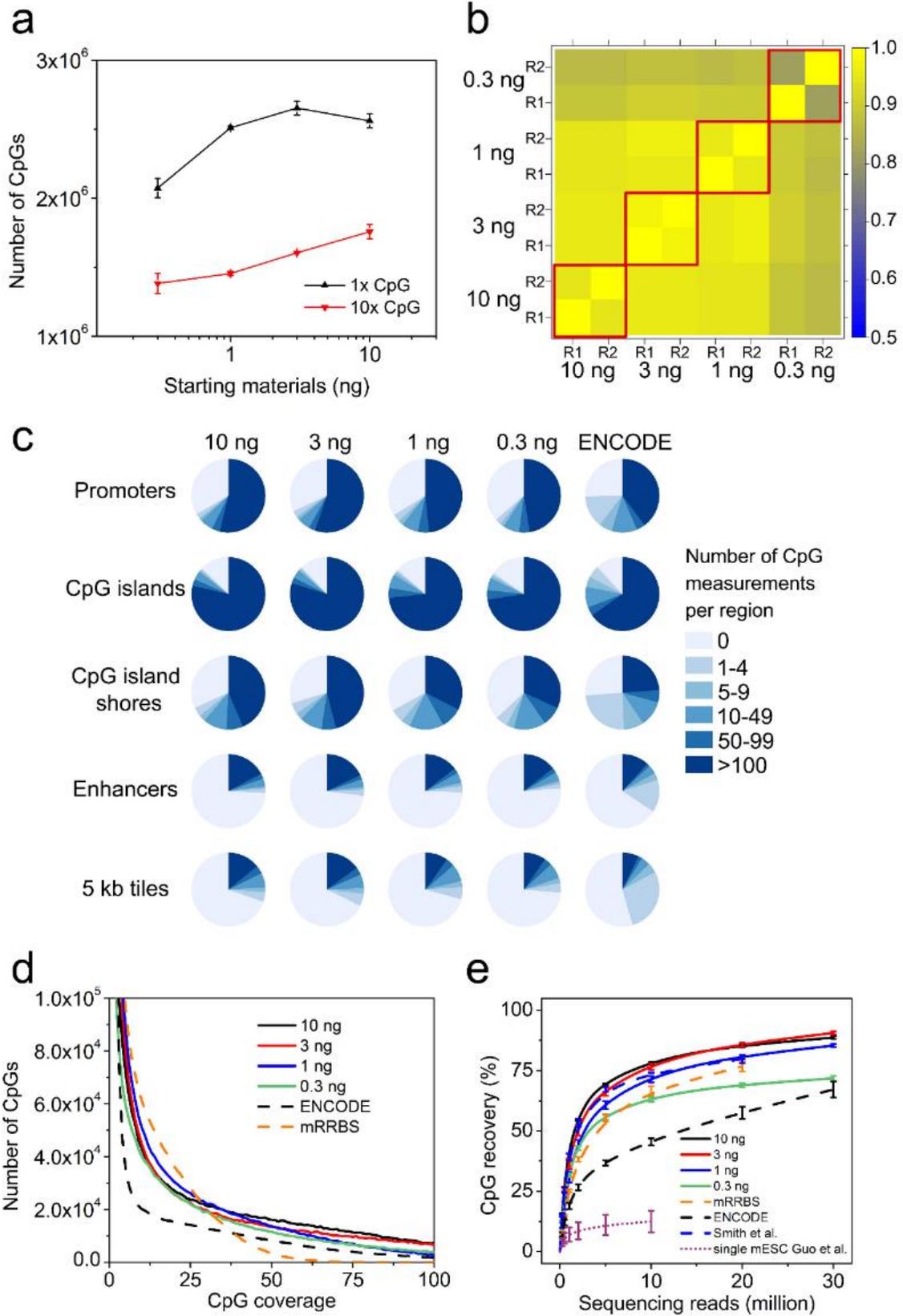
Fig. 4.4 MID-RRBS generated high quality data using sub-1 ng DNA. (a) The numbers of identified CpGs at 1× and 10× coverage with starting DNA samples of various amounts. Error bars represent s.d. (b) Cross correlations in the CG methylation level among various samples. CpGs with $\geqslant 25\times$ coverage were examined in the calculations. (c) MID-RRBS coverage of gene promoters (2 kb regions upstream of transcription starting sites of RefSeq genes), CpG islands (UCSC annotation database), CpG island shores (2 kb regions adjacent to CpG islands), enhancers (regions defined by H3K4me1+ H3K27ac based on ENCODE ChIP-seq data of GM12878 cells, ENCFF001SUE and ENCFF660QDF), and 5 kb tiles (non-overlapping consecutive 5 kb windows), in comparison to those of ENCODE data (GSM683927). (d) The number of CpGs with various coverages (0-100×), in comparison to those of ENCODE data and mRRBS (Boyle et al. 2012 20). (e) Saturation analysis of MID-RRBS data in comparison with other works (Smith et al. 21 and single cell RRBS performed by Guo et al. 13). Each point on the saturation curves (four replicates of subsamplings per dataset with each curve constructed with at least two data sets. Error bars representing s.d.) was generated by random selection of a number of raw reads followed by using the same pipeline to identify unique CpGs. Percentage CpG recovery represents the number of discovered CpGs divided by the total number of CpGs subject to RRBS examination, 2.1 million for mouse or 2.8 million for human).
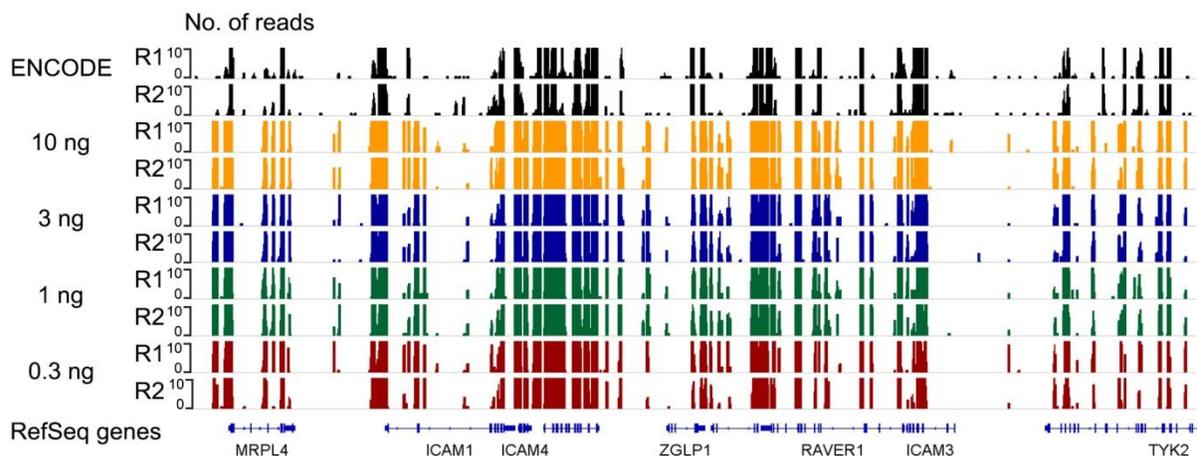
Fig. 4.5 Aligned reads (GM 12878, signal intensity up to 10× coverage) show that MID-RRBS libraries (with 0.3-10 ng input DNA) had higher complexity than those of ENCODE data (with 1 µg input DNA).

Genome-wide DNA methylation patterns in neurons and glia from mouse cerebellum are cell type-specific

We demonstrated the utility of MID-RRBS for studying primary cell samples by examining DNA methylomes in neurons and glia isolated from mouse cerebellums. Methylomic landscapes are important molecular features for defining cellular identities. Neurons and glia from human and mouse frontal cortex were previously profiled by WGBS (MethylC-seq), revealing significant cell-type specific methylomic features and associated gene activities [21, 159]. In contrast, previous reports on methylomes of mouse cerebellums only included profiles on the mixed nuclei population without separate methylomes on neurons and glia, presumably due to the limited DNA amount generated by these subpopulations [170, 171].

a

Cerebellum

Anti-NeuN staining

FACS

NeuN−
(Glia)

NeuN+
(Neuron)

FSC-A (a.u.)

Alexa 488 (a.u.)

b

mCG/CG %
0   25   50

10 ng
0.5 ng

Glia   Neuron   Mix

c

10 ng
{ NeuN+ { R1, R2 }
  NeuN− { R1, R2 } }

0.5 ng
{ NeuN+ { R1, R2 }
  NeuN− { R1, R2 } }

mCG/CG %

Rims1 mm9 chr1: 22,540,000 - 22,541,100
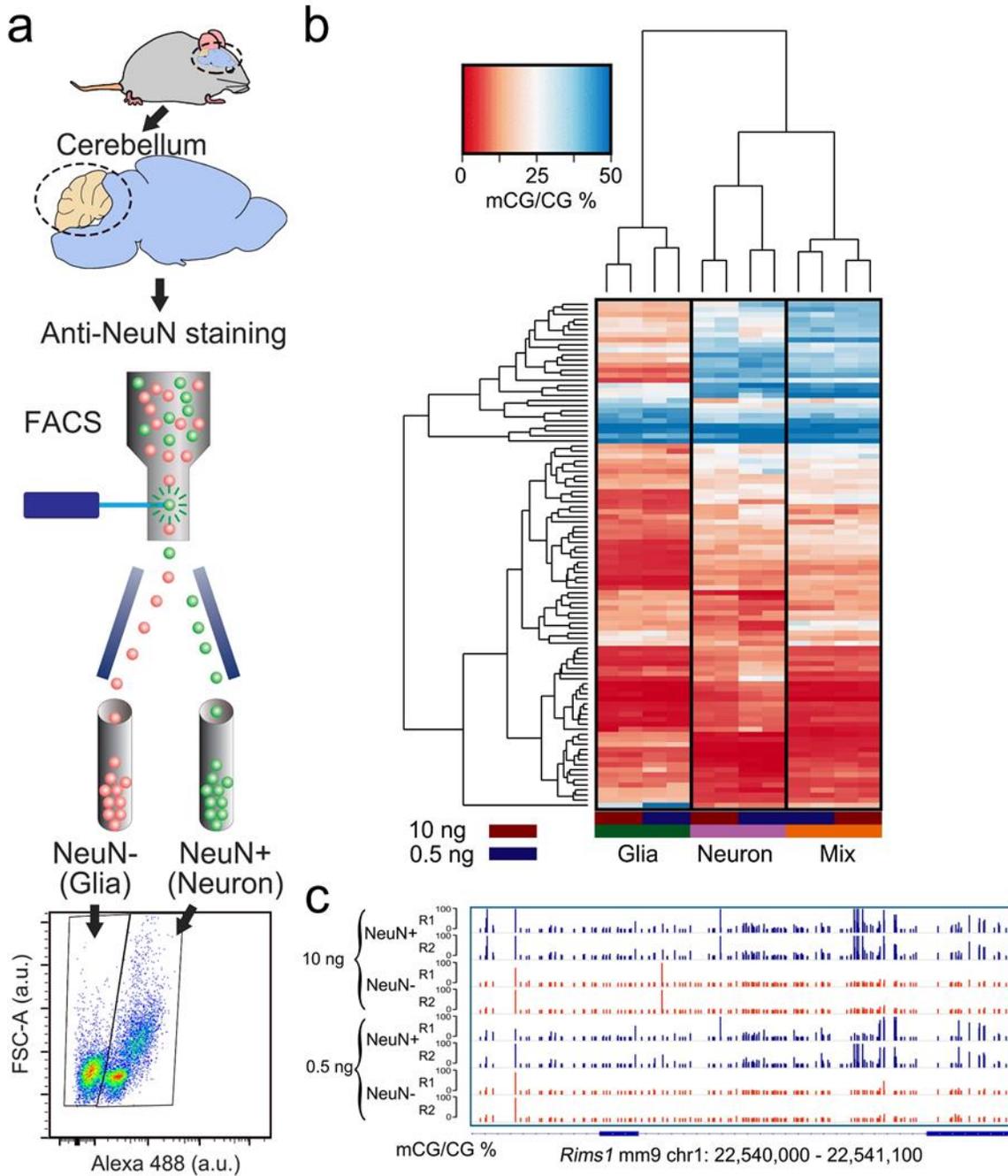
Fig. 4.6 Distinctive methylomic features in neurons (NeuN+) and glia (NeuN-) from mouse cerebellums. (a) The separation of neuronal and glial fraction from mouse cerebellum using FACS. (b) Heat map depicting unsupervised clustering analysis of 100 most variably methylated CpG islands (CG methylation) between neurons and glia. CpG islands containing more than 10 CpGs with each CpG having $\geqslant 25\times$ coverage are included in the heat map. (c) Rims1 gene CG methylation levels in neurons and glia.

In these experiments, we isolated nuclei from cerebellum samples, separated neuronal fraction (NeuN+) and glial fraction (NeuN-) by fluorescence activated cell sorting (FACS) (**Fig. 4. 6a and Fig. 4. 7**), and used MID-RRBS to examine the methylomes. We applied MID-RRBS to 10 and 0.5 ng DNA samples extracted from NeuN+ and NeuN- nuclei and achieved bisulfite conversion rate of 99.7 ± 0.1% (**Table 4. 2**). Our data on the mixed cell population ("mix") correlated well with the RRBS data in the literature (with an average Pearson correlation of 0.90). Pearson correlation coefficients among replicates of RRBS data on individual cell types were 0.98, 0.99, 0.97, and 0.98 for 10 ng NeuN+, 10 ng NeuN-, 0.5 ng NeuN+, and 0.5 ng NeuN- samples, respectively (**Fig. 4. 8**). The numbers of CpGs discovered in these samples were slightly lower than those of GM12878 samples of similar quantities (**Table 4. 1**). Our RRBS data revealed that DNA methylation was more prevalent in glia (NeuN-) than in neurons (NeuN+) in CG context (3.9% vs 2.9% by 10 ng data or 4.1 vs 2.6 by 0.5 ng data for glia and neurons, respectively, $p < 0.05$), but similar in CH context (0.16 % vs 0.18% by 10 ng data and 0.21 vs 0.23 by
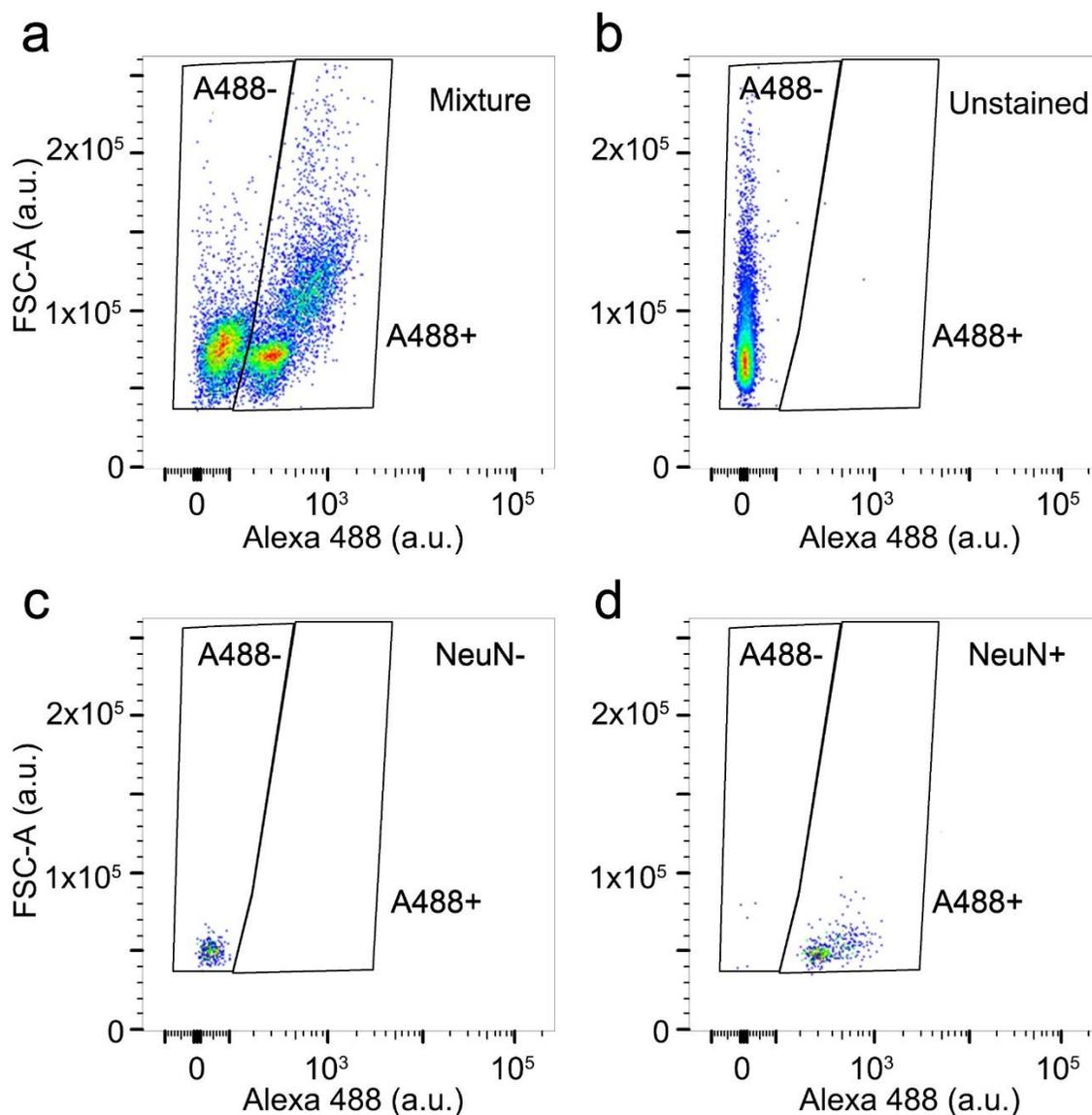
Fig. 4.7 FACS of nuclei extracted from mouse cerebellum to generate NeuN+ and NeuN- fractions. (a) Two distinct subpopulations of nuclei after staining with anti-NeuN (Alexa 488 labeled). (b) Unstained control sample. (c) NeuN- fraction after FACS sorting. (d) NeuN+ fraction after FACS sorting.

0.5 ng data for glia and neurons, respectively). We examined 100 most variably methylated CpG islands across neuronal and glial methylomes using unsupervised hierarchical clustering of CpG island methylation (**Fig. 4. 6b**). There were substantial

differences between the two cell types and the methylomes taken using the mixed cell populations only reflected the average over two cell types (**Fig. 4. 6b**). For example, *Rims1* is a gene that regulates synaptic vesicle exocytosis and neurotransmitter release [172]. Its CG methylation level in gene body was substantially higher in neurons than in glia, detected with both 10 ng and 0.5 ng samples (**Fig. 4. 6c**).

We next examined DNA methylation status in several categories of annotated gene features (promoters, intergenic regions, CpG islands, and CGI shores). The average methylation levels in glia was higher those in neurons in all categories (**Fig. 4. 9a**). Promoters presented the most significant difference ($p < 10^{-15}$, paired t-test) whereas CpG islands showed no significant difference ($p > 0.05$, paired t-test). The methylation profiles are highly reproducible between two replicates for all gene features (with an average Pearson correlation of 0.98). CpG islands were also the most conservative feature across neurons and glia (Pearson coefficient up to 0.812, **Fig. 4. 9b**), followed by CGI shores (R up to 0.763).

We examined how the CpG density of the promoters affected their cell-type specificity. We classified the promoters into three categories: High CpG-density Promoters (HCPs, CpG Ratio > 0.6 and GC% > 0.55, n=11,410), Intermediate CpG-density Promoters (ICPs, 0.4 ≤ CpG Ratio ≤ 0.6, n=3,338), and Low CpG-density Promoters (LCPs, CpG Ratio < 0.4, n=3,014) [173]. Previous works revealed that CpGs in HCPs tend to be methylated whereas CpGs in LCPs tend to be unmethylated [174, 175]. HCPs are associated with ubiquitous housekeeping and developmental genes whereas LCPs are associated with tissue-specific genes [175]. Our data strongly support this notion. Our HCPs and ICPs are largely conserved across cell types (e.g. the average Pearson

correlation coefficients were 0.83 and 0.81 between neurons and glia for HCPs and ICPs, respectively) and provide weak classifications (**Fig. 4. 10**). In contrast, LCPs show strong differential DNA methylation with cell-type specificity. The Pearson correlations between neurons and glia have an average of 0.59.
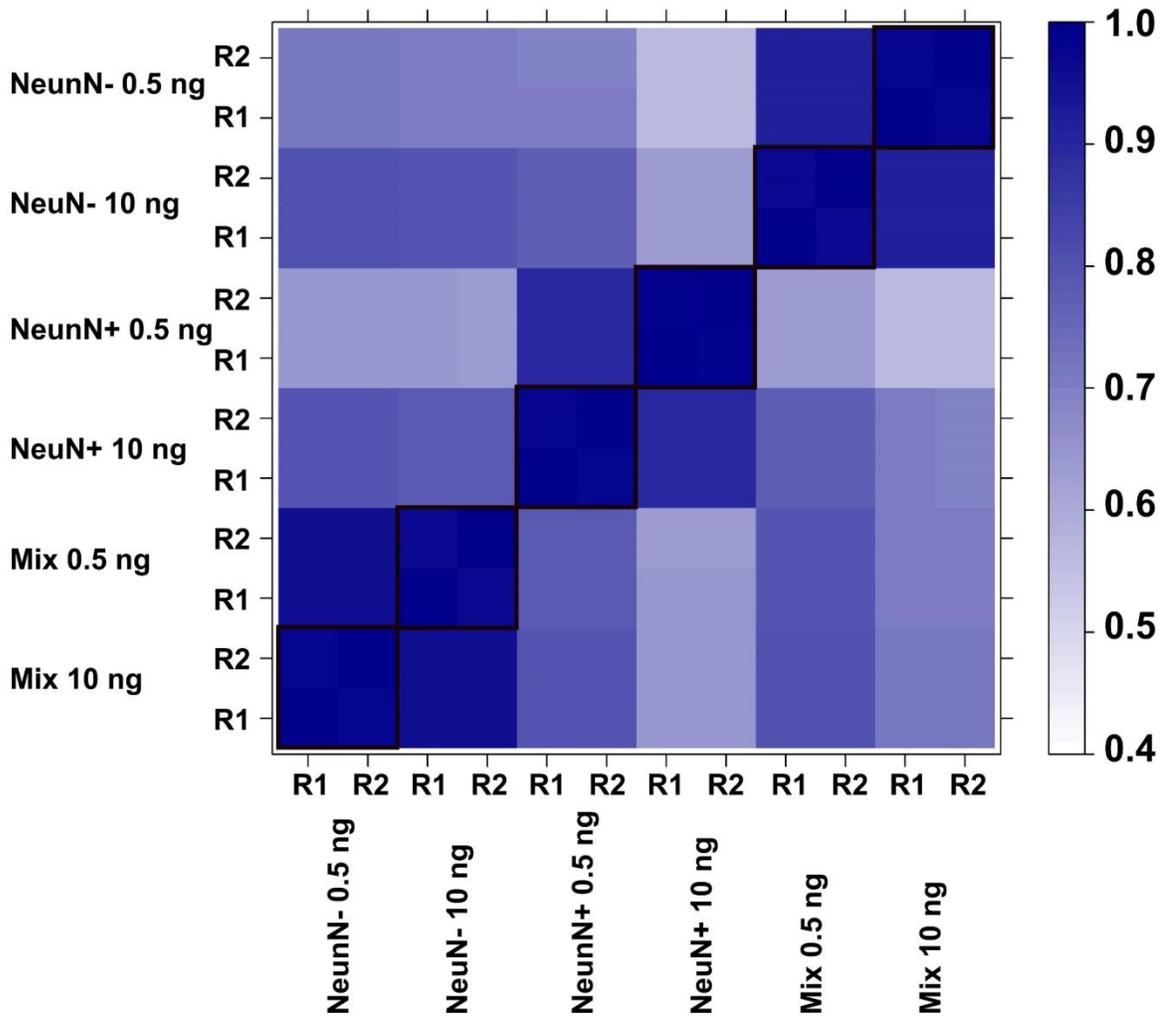
Fig. 4.8 Correlation in the CG methylation level among cerebellum (mix), NeuN+ and NeuN- samples (with CpGs having ≥25× coverage included in the calculations).
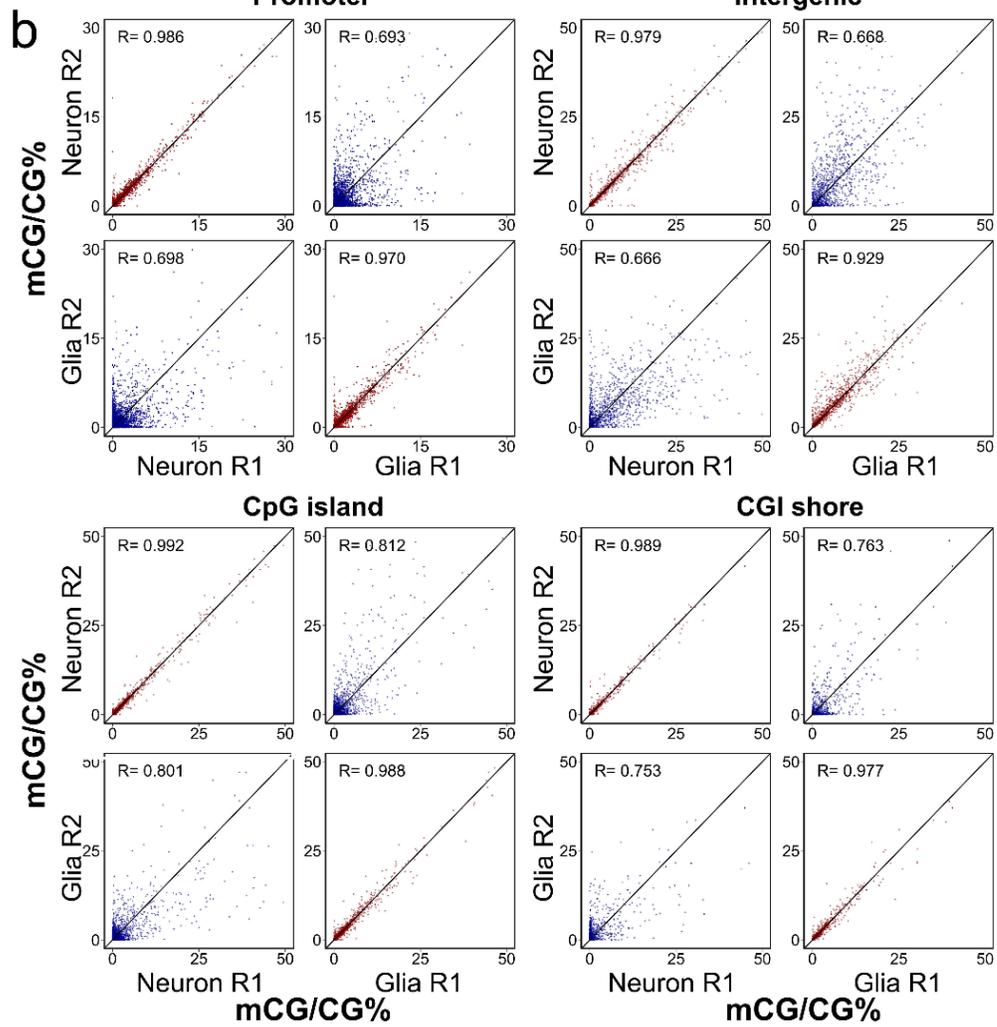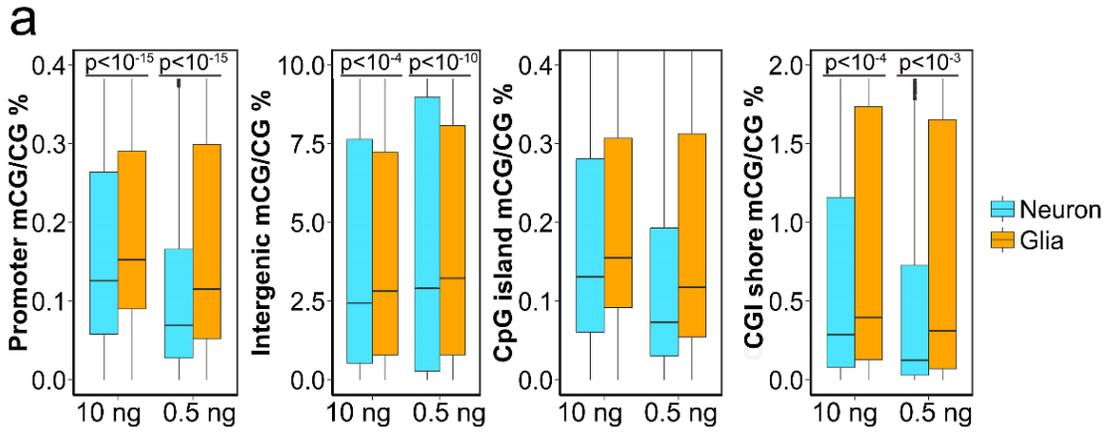
Fig. 4.9 Methylation levels across various annotated genomic features for neurons and glia from mouse cerebellum. (a) Box plots of CG methylation levels in promoters, intergenic regions, CpG islands, and CpG island shores. The boxes represent first quartile, median, and third quartile. P-values are calculated using paired t-test. (b) Scatter plots of CG methylation levels in promoters (n=23,525), intergenic regions (n=19,239), CpG islands (n=16,027), and CpG island shores (n= 30,232). The data were generated using 0.5 ng neuronal/ glial DNA. R represents Pearson coefficient. R1 and R2 are two replicates.
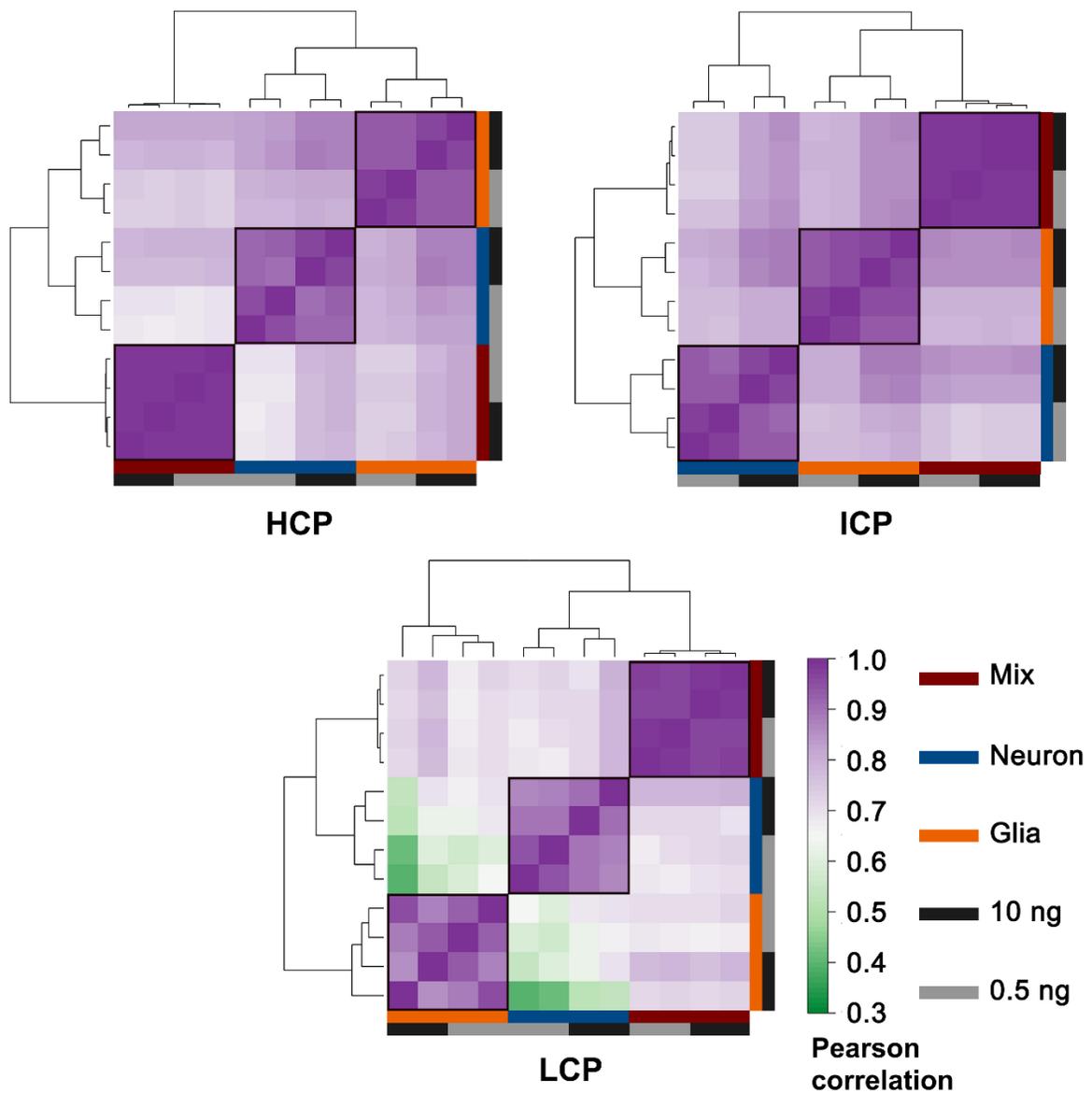
Fig. 4.10 Correlations in CG methylation levels for low, medium, and high CpG density promoters (LCP, ICP, and HCP) among cerebellum ("mix"), neuronal, and glial samples. Promoters containing more than 5 CpGs with each CpG having $\geqslant 5\times$ coverage are included in the calculations.

Differentially methylated region-associated genes are more enriched for neuron/glia-specific GO terms than differentially expressed genes

We identified 3159 and 1948 differentially CG-methylated regions (DMRs) ($p <$ 0.05) (corresponding to 2184 and 1393 DMR-associated genes), by comparing NeuN+ and NeuN- methylomes established using 10 ng and 0.5 ng samples, respectively (**Fig. 4. 11a)**. A large fraction (842) of the two sets of genes overlapped, supporting the consistency of MID-RRBS for profiling using DNA samples of different amounts. The DMRs discovered using 10 ng or 0.5 ng DNA show highly consistent distributions in various genomic features (annotated by Homer) (**Fig. 4. 11b**). The DMRs were largely located around genes (averagely 71%, within ± 1 kb of RefSeq genes). We performed unbiased gene ontology (GO) analyses on DMR-associated genes and found that these genes were strongly enriched in a number of neuron or glia-specific GO terms ($p < 0.001$, hypergeometric test). Some of these genes (e.g. Neurog2, Pax5, Foxn4, Lhx6, Wnt11, Ier2, Esr2, Grik3, Cacna1b, and Epha2) play important roles in neuronal and synaptic development and functions.

In order to examine the impact of methylomic features on transcription, we also performed transcriptome profiling (mRNA-seq) on neuronal and glial subpopulations of the nuclei (**Table 4. 3**) and identified 2331 differentially expressed genes (DE genes) between the two cell types (> 8-fold change, $p < 0.05$, t-test) (**Fig. 4. 12**). The mRNA-seq data on the mixed cell populations correlated very well with the ENCODE data (with an average Pearson correlation of 0.94, GSE93456). Only 299 (14%) of the DMR-associated genes identified using 10 ng datasets and 182 (13%) of those identified using 0.5 ng datasets exhibited differentiated expressions in the two cell types (**Fig. 4. 11a**), suggesting

that there were mechanisms other than CG methylation that strongly affected transcription. A large fraction (120) of these two sets of genes overlapped (**Fig. 4. 11a**). Consistent with previous report [176], CG methylation on non-CGI promoters is strongly associated with transcriptional repression in both cell types, whereas such correlations are less consistent with CGI promoters (**Fig. 4. 11b**).

Consistent with previous report, CG methylation on non-CGI promoters was strongly associated with transcriptional repression in both cell types, whereas such correlation was less consistent with CGI promoters (**Fig. 4. 11b**). CH methylation did not present strong correlation with the level of transcription ((**Fig. 4. 13**).
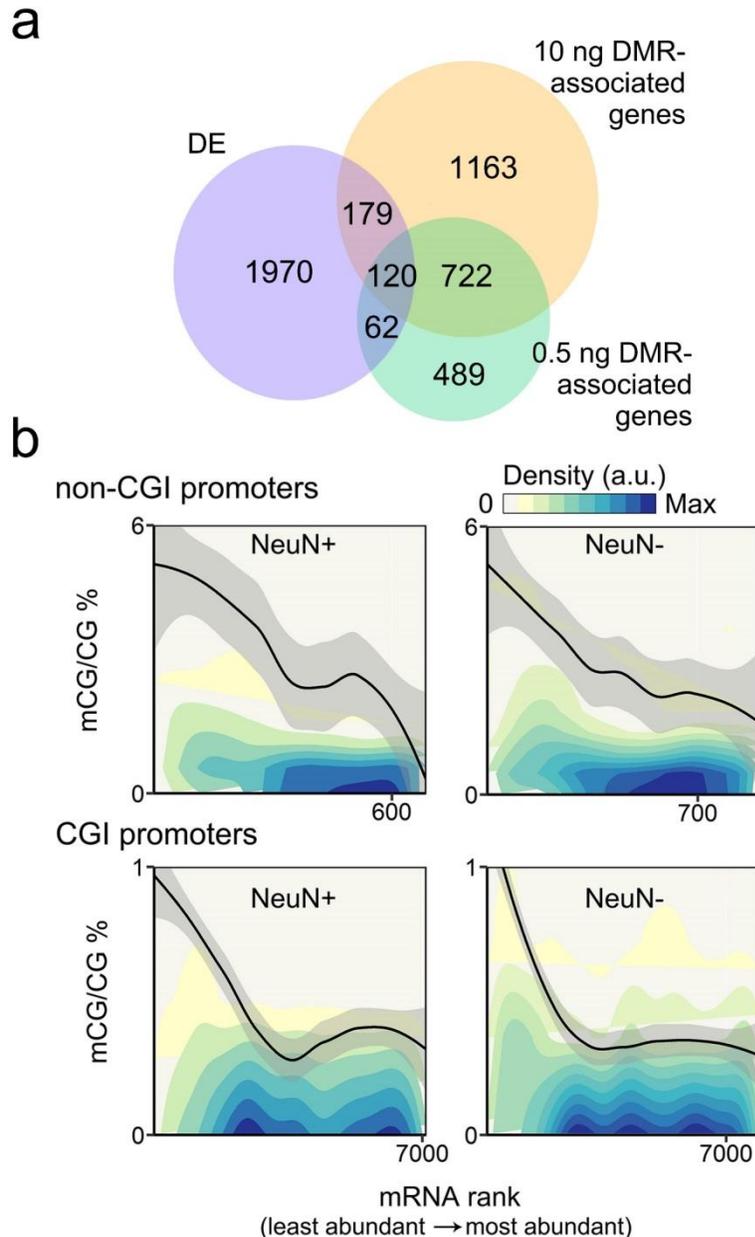
Fig. 4.11 Interactions between methylomes and transcriptomes. (a) Area-proportional Venn diagram indicating intersections among CG-DMR-associated genes discovered by experiments using 10 or 0.5 ng neuronal/glial DNA and differentially expressed genes (DE, identified by mRNA-seq). (b) The level of CG methylation in non-CGI and CGI promoters plotted against mRNA rank (generated by mRNA-seq) in neurons and glia. Grey area represents s.e.m.
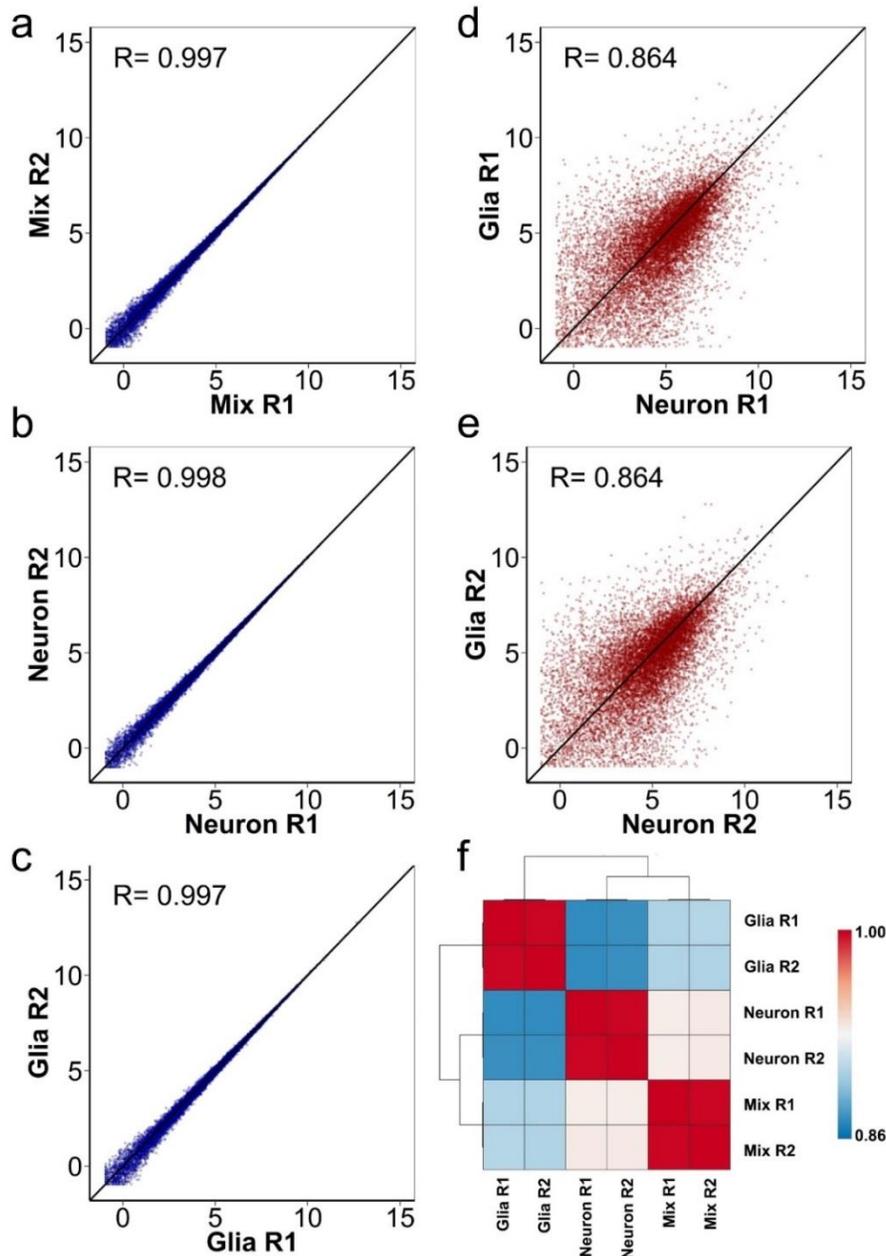
Fig. 4.12 Comparisons among mRNA-seq data of various cell types and replicates.

(a-e) Scatter plots for comparing various replicates across three cell types (cerebellum/mix, neurons and glia). (f) Correlations among cerebellum (mix), NeuN+ and NeuN- samples. All transcripts annotated in RefSeq database were used for computing correlation coefficients. The number of reads in each transcript was corrected by transcript length and total number of reads (FPKM).
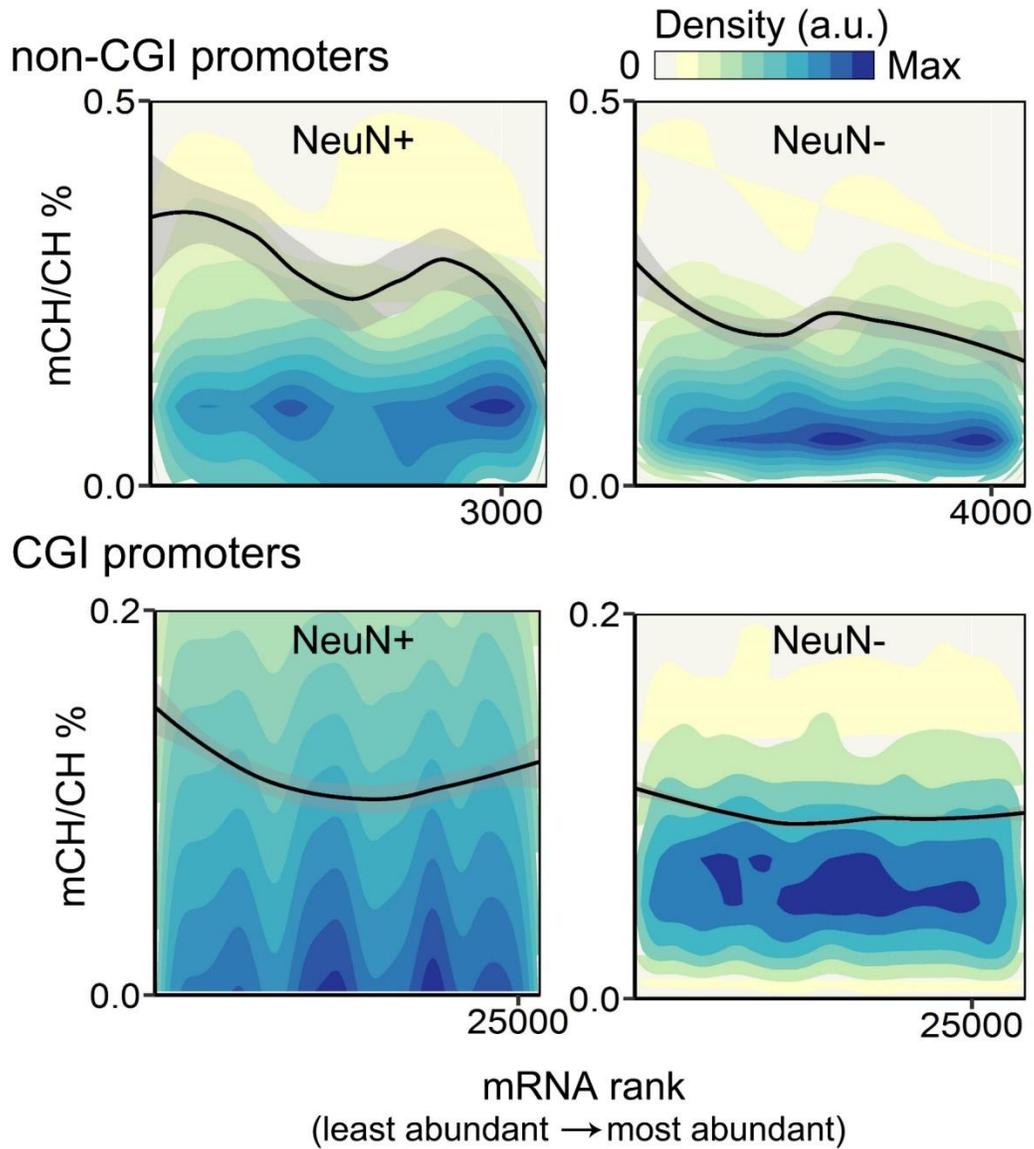
Fig. 4.13 The level of CH methylation in non-CGI and CGI promoters plotted against mRNA rank (generated by mRNA-seq) in neurons and glia. Grey area represents s.e.m.

Our MID-RRBS technology offers capability for low-input profiling of DNA methylomes on a microfluidic platform. The diffusion-based reagent exchange method permits loading/releasing of various small-molecule reagents without substantial loss in the DNA amount. Such approach facilitates conducting a complex molecular biology treatment in a microfluidic device with a simple structure. In principle, similar approaches can be applied to WGBS.

Profiling and comparing methylomes are critical for understanding difference in epigenetic regulations among various cell types in the same organ. Cell-type-specific methylome is an important part of the overall molecular machinery that defines cellular identity. In this study, we showed that our low-input MID-RRBS technology allowed us to establish distinct DNA methylomes for neurons and glia from a mouse cerebellum. The high coverage for CpGs detected by MID-RRBS ensured that various methylomic features of the two cell types were sufficiently revealed for comparison and differentiation. Complementary to previous reports on brain methylomes obtained by WGBS 4, 5, our RRBS data with focus on promoters and CGIs showed substantial difference between neuronal and glial methylomes in terms of the CG methylation levels on specific genes and across various genomic features. We found that LCPs bore high cell-type specificity. In spite of the inverse correlation between non-CGI promoter methylation and transcription within the same cell type, there was fairly little overlap between DMR-associated genes and differentially expressed genes (< 13% of DE genes and < 14% of DMR genes) when the two cell types were compared. This suggests that profiling methylomes may provide unique information for cell type classification that is orthogonal to transcriptomic analysis. We envision that the technology will generate insights into

important epigenomic dynamics using scarce primary cell samples with direct biomedical

relevance.

Table 4. 1 Summary of MID-RRBS data on GM12878 DNA samples of various starting amounts (0.3-10 ng).

Starting material (ng)	Total reads (million)	Trimmed Reads (million)	Uniquely Mapped Reads (million)	Unique Mapping efficiency (%)	1× CpG (million)	10× CpG (million)	Conversion rate (%)	Median coverage
10_R1	49.0	48.3	28.9	59.80	2.60	1.80	99.62	26
10_R2	42.6	42.1	25.3	60.00	2.52	1.72	99.61	24
3_R1	49.1	38.5	19.6	51.01	2.69	1.60	99.90	16
3_R2	39.5	35.3	20.0	56.56	2.62	1.61	99.91	17
1_R1	48.3	33.9	16.7	49.18	2.50	1.47	99.91	14
1_R2	41.4	30.1	15.3	50.78	2.52	1.44	99.91	14
0.3_R1	49.5	40.2	22.4	55.79	2.02	1.44	99.91	25
0.3_R2	36.7	29.2	15.7	53.68	2.12	1.33	99.91	17

Table 4. 2 Summary of MID-RRBS data on neuronal (NeuN+), glial (NeuN-) and cerebellum (mix) DNA samples.

Cell type/DNA amount (ng)	Total reads (million)	Trimmed Reads (million)	Uniquely Mapped Reads (million)	Unique Mapping efficiency (%)	1× CpG (million)	10× CpG (million)	Conversion rate (%)	Median coverage
Mix_10_R1	39.8	39.5	28.3	71.53	1.88	1.16	99.60	24
Mix_10_R2	43.2	42.9	30.6	71.20	1.90	1.18	99.60	25
Mix_0.5_R1	30.3	30.0	20.9	69.84	1.78	1.14	99.62	24
Mix_0.5_R2	32.9	32.6	22.7	69.57	1.79	1.16	99.62	26
NeuN+_10_R1	43.3	39.4	21.4	54.26	1.61	1.03	99.77	27
NeuN+_10_R2	44.3	40.7	23.0	56.53	1.62	1.05	99.78	29
NeuN-_0.5_R1	41.9	38.6	21.0	54.42	1.42	0.86	99.76	21
NeuN-_0.5_R2	41.9	38.9	22.1	56.73	1.43	0.87	99.78	21
NeuN+_10_R1	37.4	34.1	23.6	69.32	1.80	1.09	99.81	23
NeuN+_10_R2	38.6	35.7	25.1	70.52	1.82	1.11	99.81	24
NeuN-_0.5_R1	37.6	35.3	23.8	67.29	1.58	1.07	99.79	32
NeuN-_0.5_R2	40.0	37.9	25.9	68.35	1.59	0.94	99.80	34

Table 4. 3 Summary of RNA-seq data on neuronal (NeuN+), glial (NeuN-) and cerebellum (mix) samples.

Sample	Total reads (million)	Trimmed reads (million)	Aligned reads (million)	Alignment efficiency (%)
Mix_R1	31.2	31.2	25.8	82.73
Mix_R2	26.7	26.7	22.4	83.65
NeuN+_R1	31.3	31.3	26.0	83.22
NeuN+_R2	30.7	30.7	25.6	83.43
NeuN-_R1	32.1	32.1	24.9	77.53
NeuN-_R2	27.5	27.5	21.2	77.16

4.3 Materials and Methods

Fabrication of the microfluidic device. The polydimethylsiloxane (PDMS)-glass device was fabricated by multilayer soft lithography [29, 177]. Three photomasks (1 for the control layer and 2 for the fluidic layer) were generated with the microscale patterns designed using LayoutEditor (juspertor GmbH) and printed on high-resolution (10,000 d.p.i.) transparencies. The control layer master was fabricated by spinning SU-8 2025 (Microchem) on a 3-inch silicon wafer (P(100), 380 μm thickness, University Wafers) at 500 rpm for 10 s followed by at 3000 rpm for 45 s (yielding 24 μm in SU-8 2025 thickness). The fluidic layer master was fabricated in SU-8 2025 and AZ 9260 (Clariant) on another silicon wafer with the thickness being 60 and 13 μm, respectively. The fluidic layer master was baked at 130 ℃ for 30 s to round AZ 9260 features so that the resulted fluidic channels could be fully closed by microscale pneumatic valves. The control layer (~0.4 cm thick) was fabricated by pouring PDMS prepolymer (General Electric silicone RTV 615, MG chemicals) with a mass ratio of A: B =5: 1 onto the control layer master in a petri dish. The fluidic layer (~80 μm thick) was fabricated by spinning PDMS (A: B=20: 1) at 500 rpm for 10 s and then at 1500 rpm for 30 s. Both layers were baked at 75 ℃ for 15 min. The control layer was then peeled off from the master and access holes to the control layer channels were punched. The two PDMS layers were aligned, brought into contact, and baked for 1 h at 75 ℃ for thermal boning. The bonded PDMS structure was peeled off from the fluidic layer master and access holes to the fluidic layer were punched. The structure was then bonded to a cover glass (VWR) that was pre-cleaned in 27% NH_4OH: 30% H_2O_2: H_2O=1: 1: 5 (volumetric ratio) at 80 ℃ for 1 h. Both PDMS surface and glass

surface were oxidized in plasma (PDC-32G, Harrick Plasma) for 1 min and then immediately brought into contact. The entire structure was then baked at 75 ℃ for 1 h to strengthen the bonding.

Cell culture. GM12878 cells were purchased from Coriel Institute for Medical Research and cultured in RPMI 1640 medium (11875-093, Gibco) plus 15% fetal bovine serum (26140-079, Gibco) and 100 U/ml penicillin-streptomycin (15140-122, Gibco) at 37 ℃ in a humidified incubator with 5% CO_2. Cells were sub-cultured every 2-3 days to maintain them in exponential growth phase.

Mouse strain and cerebellum dissection. The mouse line C57BL/6 was bred and maintained in groups of 3-4 per cage with food ad libitum. 8-10 week old mice were anesthetized with isoflurane and decapitated. The cerebellum was rapidly dissected and frozen on dry ice and stored at -80 ℃ until used for nuclei isolation.

Nuclei isolation from mouse cerebellum. We used a nuclei isolation protocol that was similar to previous report [178]. A cerebellum was placed in 5 ml ice-cold nuclei extraction buffer (0.32 M sucrose, 5 mM $CaCl_2$, 3 mM $Mg(Ac)_2$, 0.1 mM EDTA, 10 mM Tris-HCl and 0.1% Triton X-100) with freshly added 50 µl protease inhibitor cocktail (P8340, Sigma-Aldrich), 5 µl of 100 mM PMSF and 5 µl of 1 M DTT. In addition, 7.5 µl of 40 U/µl RNase inhibitor (N2611, Promega) was also added when RNA-seq was conducted. Once thawed, the tissue was homogenized by slowly doucing for 15 times with a loose pestle (D9063, Sigma-Aldrich) and 25 times with a tight pestle (D9063, Sigma-Aldrich). The homogenate

was filtered with 40 μm nylon mesh cell strainer (22363547, Fisher Scientific) and then transferred into a 15 ml centrifugation tube. The sample was centrifuged at 1000 g for 10 min. The supernatant was removed and the cell pellet was gently resuspended in 1 ml cold nuclei extraction buffer with freshly added 10 μl protease inhibitor cocktail, 1 μl of 100 mM PMSF and 1 μl of 1 M DTT. 1.5 μl of 40 U/μl RNase inhibitor was mixed with the nuclei extraction buffer when RNA-seq was conducted. The 1 ml sample suspension was split into two 1.5-ml micro-centrifuge tubes (500 μl in each tube). Each sample (500 μl) was gently mixed with 0.75 ml of 50% iodixanol to yield 30% iodixanol solution. 50% iodixanol was prepared by adding 0.4 ml diluent (150 mM KCl, 30 mM MgCl2, and 120 mM Tris-HCl, pH 7.8) to 2 ml of 60% iodixanol (D1556, Sigma). The samples were then centrifuged at 10,000 g for 20 min before the supernatant was removed. The nuclei in each tube were incubated on ice for 10 min in 0.5 ml Dulbecco's phosphate-buffered saline (14190144, Life Technologies) containing 2.0% normal goat serum (50062Z, Life Technologies). The nuclei were resuspended and then pooled together (generating ~1 ml in total). For RNA-seq, 3 μl of 40 U/μl RNase inhibitor was added into the nuclei suspension. The integrity and the number of nuclei were checked under the microscope. 32 μl of 2 ng/μl anti-NeuN antibody conjugated with Alexa 488 (MAB377X, EMD Millipore) was incubated with the nuclei suspension (~1 ml) for 1 h at 4 ℃ on an end-to-end rotator. The stained samples were sorted by FACS (BD FACSAria, BD Biosciences) with 50,000 to 100,000 unlabeled nuclei used as non-staining control. We typically isolated 0.5 million neuronal nuclei and 0.5 million glial nuclei from a mouse cerebellum. The isolated nuclei were used for RNA and DNA extraction within 1 h after FACS to maintain transcriptomic and epigenomic state of nucleic acids.

Genomic DNA extraction. Human genomic DNA was purified from GM 12878 cells using QIAamp DNA Blood Mini Kit (Qiagen) following the manufacturer's protocol and suspended in EB buffer (Qiagen). Mouse genomic DNA was purified using QIAamp DNA Blood Midi Kit (Qiagen) from nuclei and suspended in AE buffer (10 mM Tris-Cl, 0.5 mM EDTA, pH 9.0). The eluate was reloaded on the column membrane and spun to maximize DNA yield. Mouse DNA was concentrated by ethanol precipitation and suspended in EB buffer. The extracted DNA was used for RRBS library construction without storage.

RRBS library construction. Library construction was conducted based on the protocol published by Gu et al. [162] and modified for Illunima Hi-Seq system. 5 µl genomic DNA was digested by adding 1 µl of 20 U/µl MspI enzyme (R0106S, NEB), 2 µl NEB buffer2 and 12 µl H_2O at 37 ℃ for 1 h and inactivated at 80 ℃ for 20 min. The fragment size was examined using 2.5% agarose gel (16550100, Thermo Fisher) in 0.5× TBE buffer (15581044, Thermo Fisher). The 3' terminal of digested DNA was end-repaired and an extra A (required by adapter ligation) was added on both strands in a single reaction. 1 µl of 5 U/µl Klenow exo- enzyme (M0212M, NEB) and 1.1 µl dNTP mixture (10 mM dATP, 1 mM dGTP and 1 mM dCTP, N0446S, NEB) were added into 20 µl digested sample. Ten-fold excess of dATP was added to increase A-tailing efficiency. The sample was incubated at 30 ℃ for 20 min (end-repair), 37 ℃ for 20 min (A-tailing) and 75 ℃ for 20 min (enzyme inactivation). The DNA (22.1 µl) was then purified by ethanol precipitation. DNA pellet was then resuspended in 17 µl EB buffer and transferred to a new 200 µl PCR tube. 1 µl of 2000 U/µl T4 ligase (M0202T, NEB), 2 µl of supplied 10× T4 ligase buffer,

and 0.6 µl NEXTflex bisulfite-seq barcode (1.6-6 µM) were mixed with the sample for ligation. The ligation reaction was carried out at 16 ℃ for 16 h and followed by 65 ℃ for 10 min for heat inactivation. The heating of the thermal cycler lid was turned off to protect the T4 ligase. Gel-free procedure [168, 179] instead of gel cutting was used for RRBS library construction. Ligated DNA (20 µl) was purified using AMPure XP beads with bead suspension to sample volumetric ratio of 1.5: 1. The DNA was eluted from Ampure beads by 20 µl EB buffer. The eluate was concentrated by ethanol precipitation and then suspended in the EB buffer for loading into the microfluidic device.

Operation of the microfluidic device. The control layer channels were filled with DI water before experiments. The on-chip pneumatic valves were actuated at 25 psi by solenoid valves (18801003-12V, ASCO Scientific) that were connected to a compressed air outlet and operated by a LabVIEW program via a computer and a data acquisition card (PCI-6509, National Instruments).

DNA (after adapter ligation and size selection) and reagents were delivered into the microfluidic device by a syringe pump. A syringe and a connected tubing was initially filled with water. An air plug (~ 1 cm long) was created (by aspiration) in the tubing to separate water from a liquid plug of DNA sample or other reagents (aspirated from a microcentrifuge tube after the air plug formation). Our microfluidic bisulfite conversion involved several steps: *1).* A conversion buffer from EZ DNA methylation-lightning kit (D5030-1, Zymo) with added 1.6% Tween 20 was loaded into the syringe pump and delivered into the two loading chambers and reaction chamber from the reagent inlet. Then the valves between the loading and reaction chambers were closed and the DNA

sample (with concentrations ranging roughly from 1.1 to 36.2 ng/µl) was loaded into the reaction chamber (240 nl) from the sample inlet, replacing the conversion reagents in the reaction chamber. *2).* The pneumatic valves were opened for 10 min to allow diffusion-based reagent exchange. The pneumatic valves were then closed. The bisulfite mix in the loading chambers was refreshed before the pneumatic valves were opened for another 10 min. This loading process was repeated one more time to reach a total diffusion time of 30 min for bisulfite mix. *3).* While the pneumatic valves were closed, the bisulfite mix in the loading chambers was flushed out by water (to avoid damage to the PDMS device during heating) and the hydration line on top of the reaction chamber was pressurized (at 25 psi) to minimize water loss during heating. The microfluidic device was then placed on a flat-plate thermal cycler (TC-4000, Techne) with 1 ml paraffin oil (18512, Sigma-Aldrich) added between the device and the thermal cycler surface to improve heat transfer. The device was heated at 98°C for 5 min (heat denaturation) followed by 55 ℃ for 1 h (sulphonation). The whole process was conducted in dark to prevent degradation of bisulfite mix. *4).* Freshly-prepared 0.3 M NaOH was delivered from reagent inlet into the loading chambers. Diffusion-based loading of NaOH (and releasing of bisulfite mix) was conducted for a total of 30 min while the NaOH solution in the loading chambers was refreshed after each 10-min period. The microfluidic device was then incubated at room temperature for 20 min for desulphonation (with the pneumatic valves closed). *5).* Converted DNA was eluted by 100 µl Tris-EDTA buffer and then purified by ethanol precipitation with the following steps. 60 µl of 5 M ammonia acetate, 4 µl of 5 µg/µl glycogen and 480 µl of ice-cold 100% ethanol were mixed with the eluate and incubated at -80 ℃ for 1 h. The mixture was then spun at 16,100 g for 15 min. The supernatant was

discarded without disturbing the precipitated pellet. 500 μl of 70% ethanol was added before the sample was spun again at 16,100 g for 10 min. The supernatant was removed and the pellet was air dried for 10 min. The pellet was finally resuspended in 20 μl EB buffer and stored at -80 ℃ until use.

PCR amplification and sequencing. The converted DNA (20 μl) was amplified by PCR (via adding 1 μl of 2.5 U/μl PfuTurbo C_x Hotstart DNA Polymerase, 5 μl of 10× PfuTurbo C_x reaction buffer, 1.25 μl of 10 mM dNTP, 3 μl of 5 μM primers, 2.5 μl of 20× EvaGreen Dye, and 17.25 μl H_2O into the DNA sample). DNA was denatured at 95 ℃ for 2 min, amplified in a number of cycles (95 ℃ for 30 s, 65 ℃ for 30 s and 72 ℃ for 45 s) depending on the amount of starting DNA (13 cycles for 10 ng, 14 cycles for 3 ng, and 16-18 cycles for 1-0.3 ng). After PCR amplification, a double size selection using AMPure XP beads was performed following the product manual. Briefly, large DNA fragments were removed by adding 30 μl AMPure bead suspension and incubating for 5 min. The beads that had large DNA bound were discarded and the supernatant was preserved. 30 μl AMPure bead suspension were then added into the supernatant and incubated for 5 min. The supernatant was discarded and DNA bound to beads was eluted into 7 μl EB buffer. Library fragment size was determined using high sensitivity DNA analysis kit (Agilent) on a TapeStation. Each library was quantified by KAPA library quantification kit (Kapa Biosystems). Each library was pooled at 10 nM for sequencing by Illumina HiSeq 4000 with single-end 50 nt read. Typically, 30-40 million reads were generated per library.

RRBS data analysis. Sequencing reads were trimmed by trim galore <u>with --RRBS option</u>. The trimmed reads were aligned to hg19 or mm9 genome by bismark [180] and bowtie [181] with default settings. The bisulfite conversion rate and CpG coverage were calculated by methylKit [182] using aligned reads. The called CpGs were visualized in IGV genome browser [183]. Regional analysis was performed for mouse data using (<u>http://www.bioinformatics.bbsrc.ac.uk/projects/seqmonk/</u>). Features containing at least 10 CpGs with each CpG having 25✕ coverage are included in the calculations. The promoters were defined as 2 kb upstream and 500 bp downstream of transcription starting sites of RefSeq genes. Intergenic regions were defined by regions other than gene bodies and promoters. CpG islands were download from UCSC database. CpG shores were defined as 2 kb regions flanking CpG Islands. The promoters were classified into HCPs, ICPs, and LCPs based on CpG ratio and GC% content. The coordinates of classified promoters were obtained from literature [173]. DMRs were identified by DSS [184] ($p < 0.05$) and annotated to the closest RefSeq genes (i.e. DMR-associated genes). Genes that were closest to DMRs were extracted as targets for gene ontology (GO) term enrichment analysis by DAVID [185] ($p < 0.05$).

mRNA-seq. RNA was extracted from ~100,000 nuclei using RNeasy Mini Kit (Qiagen). DNase treatment was included following manufacturer instructions to remove gDNA contamination. The RNA quality was detected by High Sensitivity RNA ScreenTape System (Agilent) to yield RNA integrity number (RIN) > 6. mRNA-seq libraries were prepared using SMART-seq v4 Ultra Low Input RNA kit (Clontech) and Nextera XT DNA library prep kit (Illumina) following instructions. Briefly, 1 ng RNA (9.5 µl) was mixed with

1 µl of 10× reaction buffer. The sample was mixed with 2 µl of 3' SMART-seq CDS primer II A and incubated at 72 ℃ for 3 min. The sample was then placed on ice for 2 min. 5.5 µl Master Mix and 2 µl SMARTScribe Reverse Transcriptase were added into the sample (20 µl in total). The mixture was incubated at 42 ℃ for 90 min and then 70 ℃ for 10 min to synthesize first-strand cDNA. cDNA was then amplified by 13 cycles of PCR. The PCR product was purified by AMPure beads. 100-150 pg purified cDNA was used for Nextera library preparation. The cDNA was tagmented at 55 ℃ for 5 min and neutralized at room temperature for 5 min. The tagmented DNA was amplified by 12 cycles of PCR and purified by 30 µl AMPure XP beads. The fragment size was determined by a Tapestation and quantified by KAPA qPCR library quantification kit. Each library was pooled at 10 nM for sequencing by Illumina HiSeq 4000 with single-end 50 nt read. Typically, 20 million reads were generated per library.

mRNA-seq data analysis. Sequencing reads were trimmed by trim galore with default settings. The trimmed reads were aligned by Tophat [186]. The aligned reads were further analyzed by cufflinks [187]. The gene expression level was plotted by cummerbund (http://compbio.mit.edu/cummeRbund/), Seqmonk and custom R scripts. Differentially expressed genes were identified by cummerbund and used as targets for gene ontology (GO) term enrichment analysis by DAVID ($p < 0.05$).

Mathematical modelling of microfluidic diffusion. The time-dependent diffusion process in the microfluidic device was modelled using COMSOL Multiphysics 4.3 as previously

described [9]. We used $1.3×10^{-9}$, $9×10^{-10}$ and $2×10^{-11}$ for diffusivities of ions, quinol and 250 bp DNA, respectively [188-190].

CHAPTER 5 – Multiplexed Microfluidic Device for High-throughput and Low-input Histone Modification Analysis Based on Surface Immobilized antibody for ChIP-seq (SurfaceChIP-seq)

5.1 Introduction

Protein-DNA interaction and histone modifications play critical roles in brain function and development by regulating gene expression. Distinguishable epigenomic landscape exists in seemingly homogeneous tissue/ cell population (e.g. cerebellum). Chromatin immunoprecipitation coupled with next generation sequencing (ChIP-Seq) is the gold standard for studying *in vivo* genome-wide protein-DNA interactions and chromatin modifications. ChIP-seq typically involves cell lysis, chromatin fragmentation (e.g. enzyme digestion or sonication), immunoprecipitation of targeted chromatin using magnetic beads, digestion/removal of proteins, and amplification/identification of ChIPed DNA. There are two major limitations of the conventional ChIP-seq protocols: *1)* requirement of large number of cells (>10^7 cells) which limits application for scarce samples, *2)* tedious process with long hands-on time which compromises the reproducibility and throughput of the assay. Various strategies have been applied to improve the sensitivity of conventional ChIP-seq, such as LinDA (10000 cells) [39], nano-ChIP-seq (5000 cells) [38], iChIP (500 cells) [40], MOWChIP (100 cells) [29] and Drop-ChIP (single cells) [31]. MOWChIP uses packed magnetic beads in a microfluidic channel to enrich chromatin and oscillatory washing to reduce non-specific adsorbed proteins. It profiled histone 3 lysine 4 tri-methylation (H3K4me3) mark and histone 3 lysine 27 acetylation (H3K27ac) using as few as 100 cells. Drop-ChIP barcodes chromatin from

each cell and performs immunoprecipitation on the pooled sample with the presence of significant amount of 'carrier' chromatin. It collects chromatin data at single-cell resolution with limited genomic coverage (1000 unique reads per cell) and significant amount of uninformative reads (>97%). All the strategies above were developed based on antibody coated beads which have overwhelming surface area to adsorb chromatin. For limited starting materials (~100 cells), the non-specifically adsorbed chromatin occupies bead surface which leads to considerably low signal-to-noise ratio (enrichment). It is also challenging to automate bead-based approaches which limit the potential applications.

Here, we demonstrate a microfluidic technology, referred to as <u>Surface</u> <u>I</u>mmobilized Antibody for <u>Ch</u>romatin <u>I</u>mmuno<u>p</u>recipitation coupled with next generation <u>seq</u>uencing (SurfaceChIP-seq) for low-input (30-100 cells) profiling of multiple histone modification landscapes in parallel. The antibodies are covalently immobilized on the glass surface instead of magnetic beads. This strategy not only dramatically reduces surface area and non-specific adsorption, but also enables a fully automated immunoprecipitation procedure by simplifying chip structure. We employed a stepwise sample injection to maximize collection efficiency of ChIPed DNA and an oscillatory washing for removing nonspecific adsorption/trapping. The combination of flat immobilized antibody, stepwise sample injection and oscillatory is the key to extremely high yield of highly enriched DNA.

We mapped the epigenome of a cell line (GM12878), neuron and glia isolated from mouse cerebellum with multiple histone marks (H3K4me3, H3K27me3, and H3K27ac). Our protocol yield data with high quality (AUC 0.968 for 5000 cells, 0.927 for 100 cells, and 0.912-0.864 for 30 cells) and reproducibility (Pearson correlation 0.995 for 5000 cells

0.991 for 100 cells, and 0.897 for 30 cells). It permitted profiling up to 8 ChIP assays simultaneously with multiple histone marks. The SurfaceChIP-seq allowed differentiation of epigenomic landscape of neuron and glia, generating biological insights into cell-type specific features (bivalent promoter, enhancer and super-enhancer) and their relationships to gene expressions.
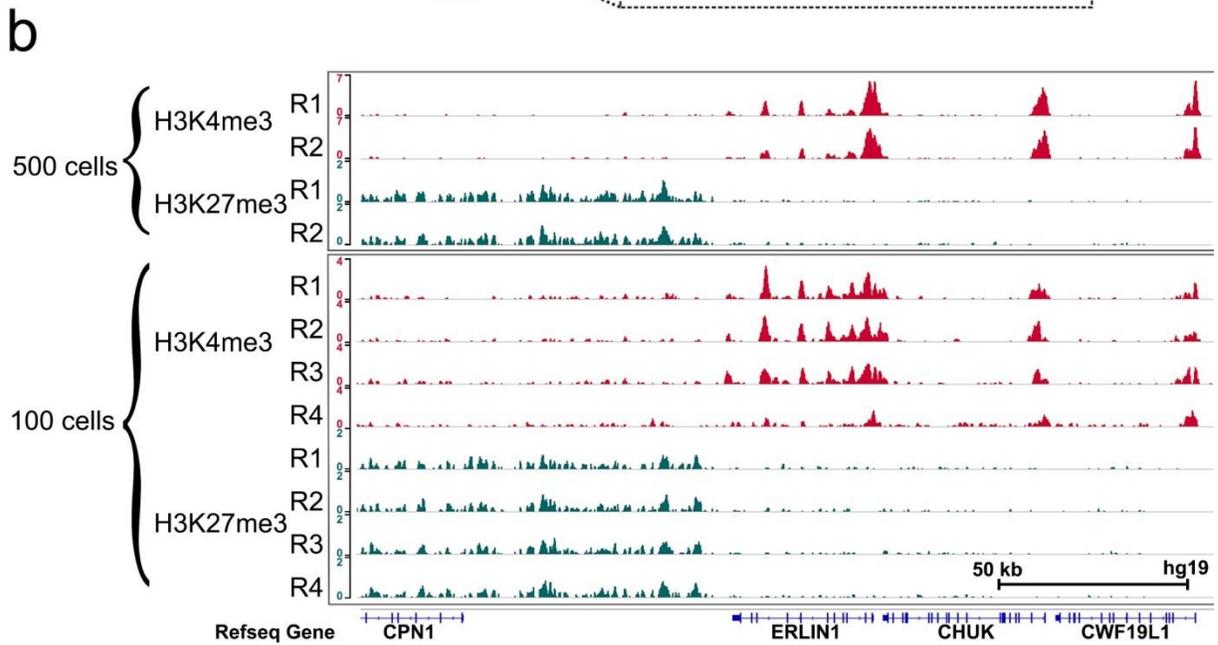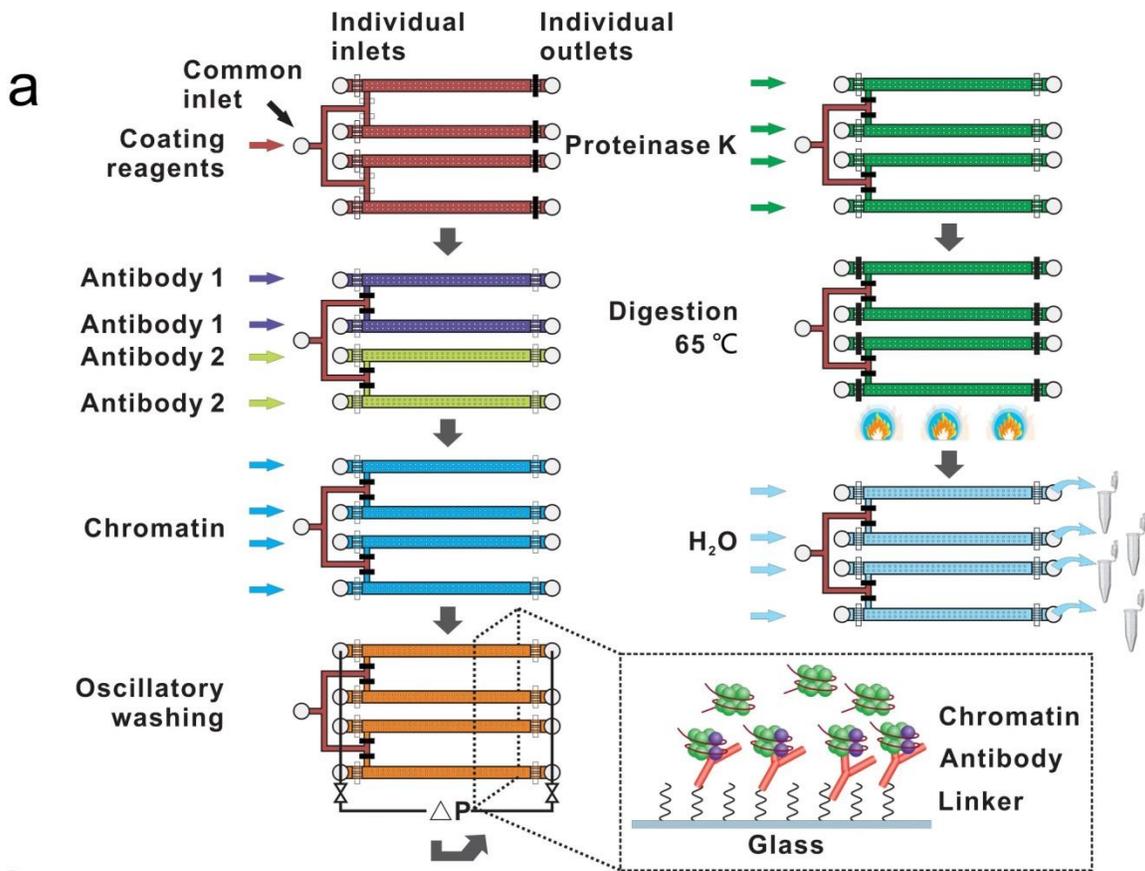
Fig. 5.1 The overview of the SurfaceChIP protocol and profiling of H3K4me3 and H3K27me3 marks in parallel using 100 or 500 cells. (a) Experimental procedure. (b) Representative normalized ChIP-seq signals of GM12878 using 100 or 500 cells in a single multiplexed device.

5.2 Results and Discussion

Multiplexed SurfaceChIP-seq for profiling multiple histone marks.

We used multi-layer soft lithography to design and fabricate a polydimethylsiloxane (PDMS) device, featuring 4 or 8 multiplexed microfluidic chambers (2.4 µl in volume, 40 mm×1 mm ×0.06 mm) for high-throughput ChIP with multiple histone marks (**Fig. 5.1a and Fig. 5.2a**). The microfluidic chambers have a common inlet, 4 or 8 individual inlets and 4 or 8 individual outlets. The inlets and outlets can be fully closed by on-chip pneumatic valves by exerting a pressure at 25 psi (**Fig. 5.3**). The common inlet allowed simultaneous surface coating of all the chambers and individual inlets permitted profiling different histone marks and/or different samples in parallel. We also fabricated a single layer PDMS device with individual microfluidic channels (40 mm×60 µm×1 mm) (**Fig. 5.2b**). This extremely simplified device allowed rapid screening of coating conditions without using any pneumatic valve during surface modification procedure. Key features and operation of the two-layered device are shown in **Fig. 5.1a** (see **Materials and Methods**). *1)* The plasma activated glass surface was covalently derived by poly-l-lysine (PLL) and DNA oligo linker. The DNA oligo increased the antibody coating density and maintained the activity of antibody [191]. *2)* DNA-conjugated antibodies were flowed into microfluidic chamber and hybridize with DNA oligo linker. *3)* Micrococcal Nuclease (MNase) digested chromatin (~170 bp) were periodically flowed through the antibody

immobilized microchambers and adsorbed on the glass surface. *4)* The microchambers were then washed by oscillatory washing in two different washing buffers (a low-salt buffer followed by a high-salt buffer) to remove nonspecifically adsorbed chromatin fragments. *5)* The captured chromatin was digested by proteinase K for 30 min to release ChIPed DNA from the surface. *6)* The ChIPed DNA was eluted for off-chip processing.

a

b

5 mm
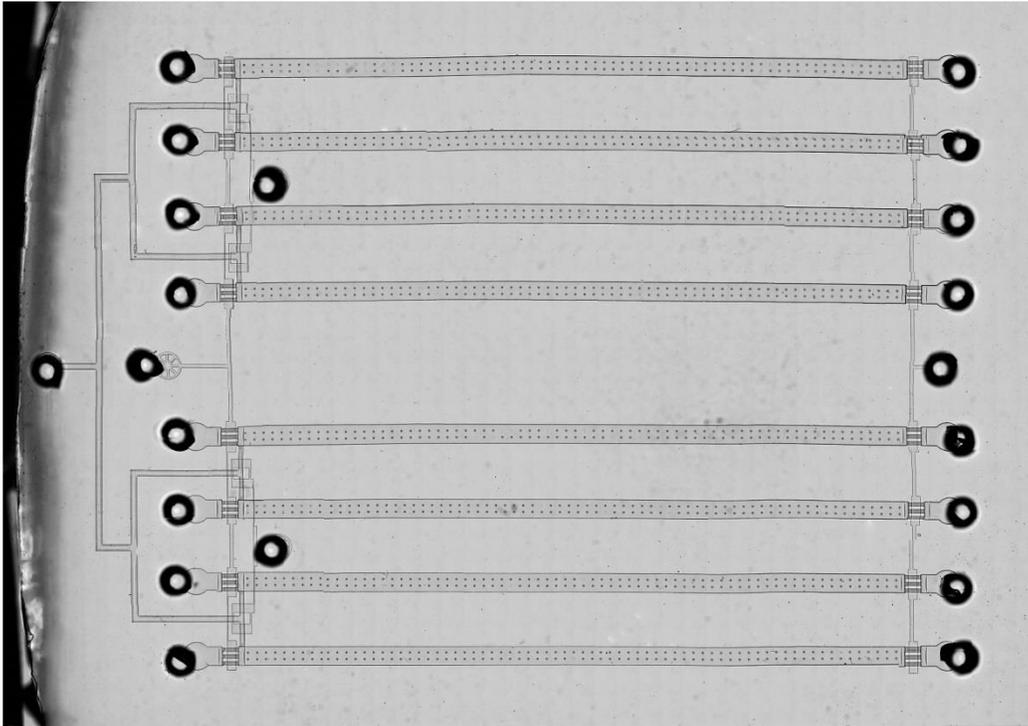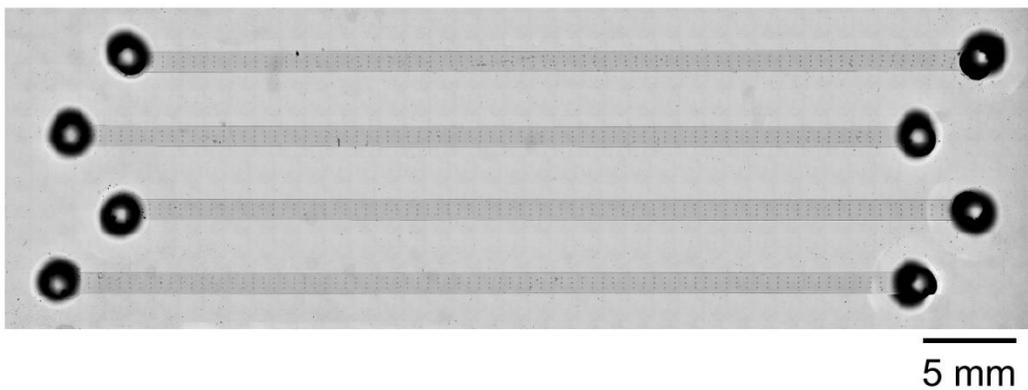
Fig. 5.2 Microscopic images of microfluidic devices. (a) 8-multiplexed multi-layer device. (b) Single-layer device.

a

b
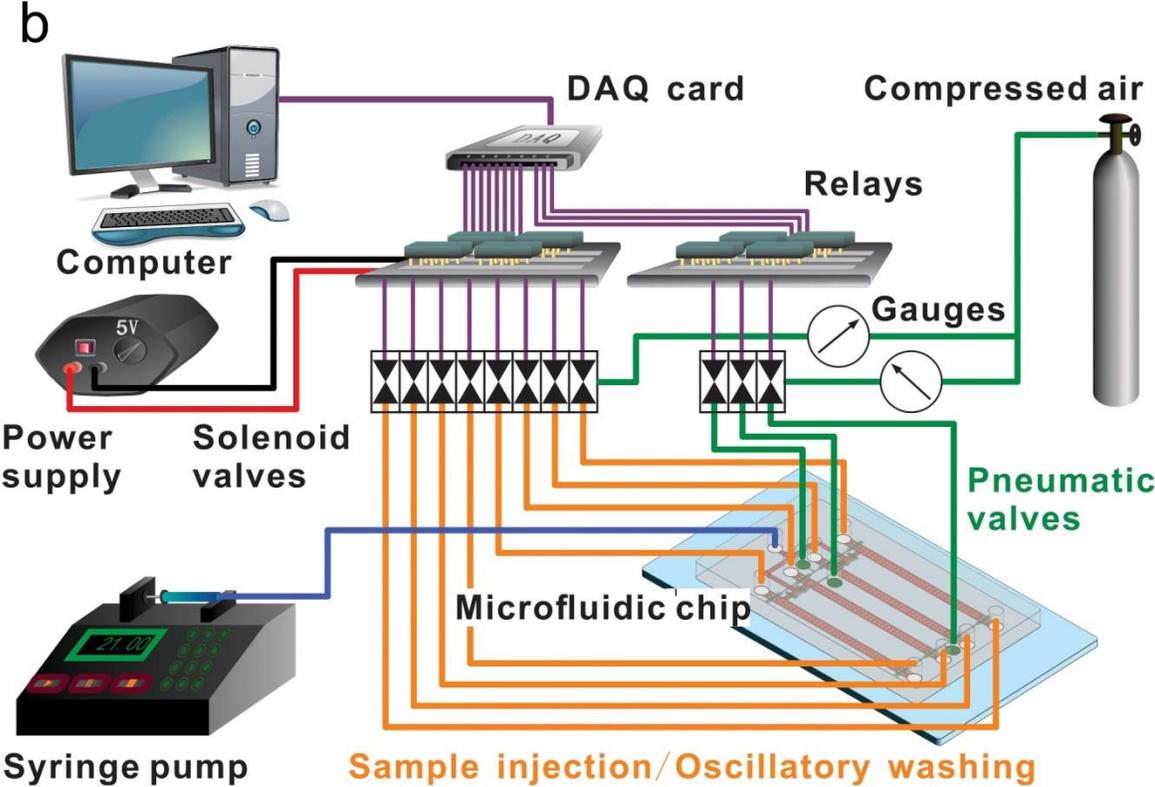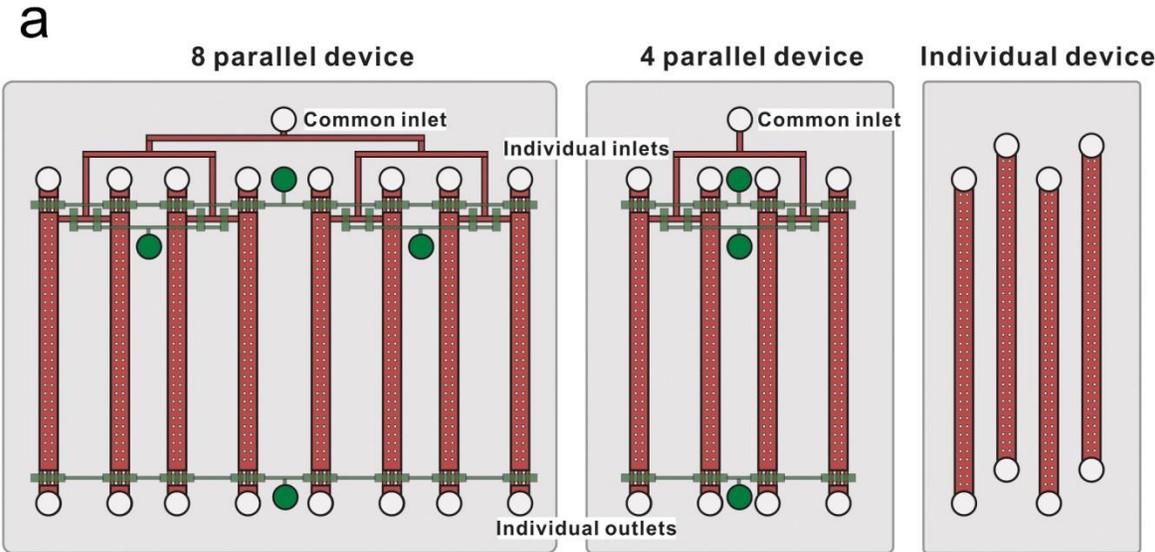
Fig. 5.3 Design and setup of the microfluidic device. (a) The design of microfluidic devices. (b) A schematic of the ancillary system for control and liquid delivery and its connection with the microfluidic device.

We found that the performance of SurfaceChIP was affected by several parameters of the protocol, including antibody concentration used for coating glass surface, duration of the oscillatory washing, duration of immunoprecipitation, and cell sample size. The careful optimization of these parameters is critical for obtaining high-quality ChIP-Seq data from tiny amount of cells. These parameters were optimized using SurfaceChIP-qPCR to evaluate the fold enrichment of known positive loci (*Zmynd11*, *Unkl*, *C9orf3,* and *Chst14*) and negative loci (N1 and N2) for histone 3 lysine 4 trimethylation (H3K4me3) in a human lymphoblastoid cell line, GM12878 (**Table 5.1**). Assuming the glass surface (40 mm × 1 mm) of the microfluidic chamber was fully covered with antibody (7-10 nm in diameter), a surface would capture ~180 ng DNA with 170 bp (mono-nucleosome) in length which is less than 5% of capture capacity (at least 4 μg DNA) in bead-based protocols. It is more than enough to perform ChIP from 10000 or fewer cells. Non-specific adsorbed chromatin will significantly reduce the signal to noise ratio (fold enrichment), thus an effective washing is essential for generating high-quality ChIP data. We found oscillatory washing was necessary to reduce nonspecific adsorption and physical trapping (**Fig. 5.4a**). Excessive washing should be avoided to prevent DNA loss. Comparing to bead-based ChIP protocols [29, 31, 38, 40, 192], non-specific chromatin is less possible to be adsorbed or trapped on the flat glass surface. The washing was also more effective due to the increased relative flow rate. We also found that the fold enrichment

of ChIPed DNA reached peak value at an intermediate antibody concentration (**Fig. 5.4b**).

The insufficient antibody coverage on the glass surface at low concentration decreases

the amount of chromatin to be captured, while excessive antibody increases the chance

that non-specific chromatin to be adsorbed. Under the optimized conditions, we obtained

~700 pg and 25 pg ChIPed DNA from 5000 and 100 for H3K4me3, respectively. The yield

was comparable to our previous work [29] but about 2 orders of magnitude higher than other

reported works [38]. We also noticed that the stepwise immunoprecipitation step improved

the capture efficiency of ChIPed DNA. Within 6 min of incubation, 99.5% of the chromatin

was adsorbed on the surface based on modelling result (**Fig. 5.5**). We replaced the

chromatin in the microfluidic chamber with fresh chromatin every 6 min to maximize the

capture efficiency and minimize non-specific adsorption. The entire sample was flowed

through the chamber in 10 6-min periods (~ 1 h in total).

Table 5.1 Real-time PCR primers.

Species	Histone mark/antibody	Name	Primer
Human	H3K4me3 EMD Millipore 07-473	*Zmynd11* (F)	AGG ATA ATC AGC CCC TGA ATA A
		Zmynd11 (R)	TCC ATC AGT CAG TCC GCA GT
		UNKL (F)	CAG CCA CCC ACC TAG GAA
		UNKL (R)	TCC TAT GGC TCC CCA GGT
		C9orf3 (F)	CCT CCT CAG TTC TCC CAG ACT
		C9orf3 (R)	AGC TGA GGT GGT AAG ATG TGA C
		Chst14 (F)	GAC CGT TAC CGC TTC CTC TA
		Chst14 (R)	CTG CCA GCA CCT TCA TCA
		N1 (F)	TCA TCT GCA AAT GGG GAC AA
		N1 (R)	AGG ACA CCC CCT CTC AAC AC
		N2 (F)	GCC CTG CCT GAA TGC TCT AG
		N2 (R)	CAA GGA CAC CCC CTC TCA AC
	H3K27me3 EMD Millipore 07-449	P2 (F)	AGT ATG TGG TCC AAA CAC TTA TCA
		P2 (R)	CCA GAC GGG TGG CAA ATA
		P4 (F)	CAG CAC CTC CAG CCC TTA T
		P4 (R)	TCC CCA GAG GGA GGA AAG
		P5 (F)	CTC GGC CTT GAT CTT CTC C
		P5 (R)	TCT GGC TTT CTT CGC TTT TC
		P6 (F)	CGT CTT CGT ATG CCA TCA AC
		P6 (R)	GTT GAA CAC GGG TCA GTC G
		N1 (F)	TAC TAG CGG TTT TAC GGG CG
		N1 (R)	TCG AAC AGG AGG AGC AGA GAG CGA
		N2 (F)	CCC ATT TCT CTC TCC ACT TCC
		N2 (R)	ACA TGC CTT AGG GTT CGT TC
		N3 (F)	TGG TAC AAG GTT GGG AGG TA
		N3 (R)	CAT CCA ATG GAG AAA AGG TAG T
		N4 (F)	ATG GTT GCC ACT GGG GAT CT
		N4 (R)	TGC CAA AGC CTA GGG GAA GA
Mouse	H3K4me3 EMD Millipore 07-473, H3K27ac Abcam Inc. ab4729	Gapdh-2	Active motif 71018
		Actb-2	Active motif 71017
		Negative-1	Active motif 71011
		Negative-2	Active motif 71012
	H3K27me3 EMD Millipore 07-449	Hoxc10	Active motif 71019
		Pax-2	Active motif 71020
		Negative-3	Active motif 71013
		Negative-4	Active motif 71014
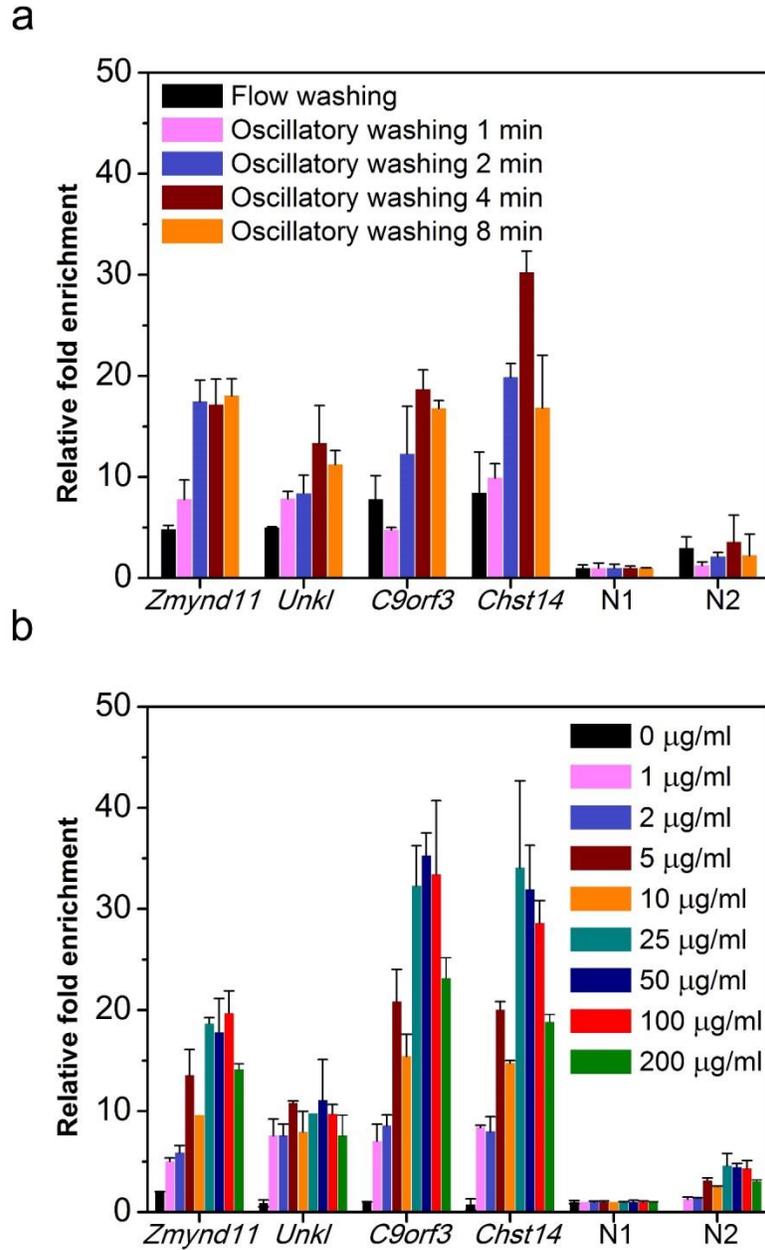
Fig. 5.4 qPCR optimization of SurfaceChIP for H3K4me3 mark from 1000 GM12878 cells. a) Optimization of washing condition b) Optimization of antibody concentration.
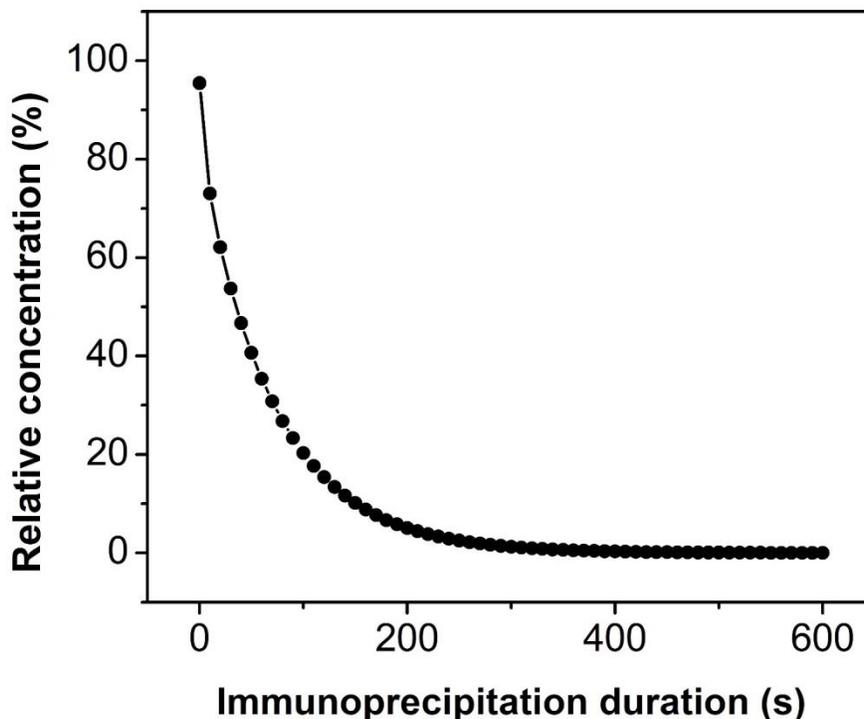
Fig. 5.5 The COMSOL modelling result of chromatin concentration change in the microfluidic chamber over time.

The SurfaceChIP-seq allows simultaneous examine multiple histone marks using as few as 100 cells for each chamber. We profiled H3K4me3 and histone 3 lysine 27 trimethylation (H3K27me3) marks simultaneously using chromatin generated from 500 or 100 GM12878 cells. The sequencing results are summarized in **Table 5.2**. **Fig. 5.1b** shows normalized signals from the 4-multiplexed device using 500 cells (Upper panel) and 8-paralleled device (lower panel) using 100 cells around *CHUK* gene which plays a critical role in the negative feedback of NF-κB signaling to limit inflammatory gene activation [193]. The distinguishable patterns between the two histone marks prove that little probability of cross contamination. Among the 4 replicates of H3K4me3 using 100 cells,

we generated 3 useable samples. The usable H3K4me3 samples of 500 and 100 cells had Pearson coefficients of 0.99 and 0.96-0.93 between replicates, respectively. The H3K27me3 samples of 500 and 100 cells had Pearson coefficients of 0.99 and 0.99-0.95 between replicates, respectively.

Table 5.2 Summary of SurfaceChIP-seq data.

Sample type	Histone mark	No. of cells	Total reads (million)	Trimmed reads (million)	Aligned reads (million)	Alignment efficiency %	Peaks
GM12878 multiplexed experiment	H3K4me3	500	28.07	28.00	27.31	97.51	46636
		500	19.96	19.79	19.10	96.54	48641
		100	20.46	20.43	19.52	95.56	33386
		100	21.11	20.87	19.74	94.55	34514
		100	14.74	14.63	13.95	95.32	43625
		100	21.15	21.14	12.25	57.94	4103
	H3K27me3	500	58.44	58.12	56.87	97.84	27643
		500	45.12	45.06	44.11	97.90	26897
		100	30.79	30.78	30.01	97.51	25841
		100	34.24	34.21	33.30	97.35	28555
		100	42.30	42.28	41.21	97.48	30541
		100	41.29	41.23	40.20	97.48	28120

Sample type	Histone mark	Total reads (million)	Trimmed reads (million)	Aligned reads (million)	Alignment efficiency %	Peaks
GM12878 ENCODE	H3K4me3 Broad	95.70	74.48	67.11	90.11	27479
		49.08	29.36	23.18	78.95	24451
	H3K4me3 UW	22.81	22.27	20.18	90.58	37562
		25.26	24.76	23.93	96.65	36413
	H3K27me3 Broad	10.30	8.66	8.20	94.70	13397
		28.75	14.97	13.85	92.52	17365
		45.95	45.63	44.55	97.64	14917
	H3K27me3 UW	19.08	18.95	18.31	96.64	16366
		14.46	14.36	13.55	94.36	17495

Sample type	Histone mark	No. of cells	Washing duration (min)	Total reads (million)	Trimmed reads (million)	Aligned reads (million)	Alignment efficiency %	Peaks
GM12878 protein A linker	H3K4me3	500	1	23.10	23.08	22.07	95.64	8
		500	1	17.88	17.87	16.85	94.29	33292
		500	2	58.78	58.63	56.05	95.60	4444
		500	2	28.26	28.14	26.82	95.32	31026
		500	4	18.79	18.54	17.65	95.20	13410
		500	4	15.49	15.27	14.44	94.52	14404
		500	8	28.77	28.55	27.30	95.61	10898
		500	8	24.55	24.46	23.31	95.32	35389
		100	4	31.05	30.61	28.77	93.99	344
		100	4	32.54	32.10	29.33	91.35	798

Sample type	Histone mark	No. of cells	Total reads (million)	Trimmed reads (million)	Aligned reads (million)	Alignment efficiency %	Peaks
GM12878 DNA oligo linker	H3K4me3	5000	20.31	20.31	19.89	97.95	47765
		5000	20.26	20.26	19.84	97.93	49853
		1000	14.99	14.99	14.68	97.95	53167
		1000	17.99	17.99	17.62	97.94	51276
		500	16.59	16.53	16.03	96.94	55333
		500	19.20	19.11	18.55	97.06	58251
		100	18.46	18.45	17.91	97.12	59602
		100	20.85	20.84	20.23	97.07	56745
		60	28.30	25.43	24.39	95.93	37705
		60	17.69	16.28	15.61	95.90	46842
		30	19.44	19.31	18.43	95.47	22673
		30	38.19	35.06	31.75	90.56	5484
		30	35.97	35.48	31.69	89.34	5396
		1 M	18.84	18.83	18.36	97.53	56659
		1 M	18.41	18.40	17.96	97.63	45249
	Input		19.45	19.38	9.94	51.31	
GM12878 DNA oligo linker	H3K27me3	10000	57.47	57.38	55.97	97.53	31466
		10000	81.79	81.67	79.64	97.51	32567
		5000	64.81	64.64	63.24	97.82	32421
		5000	57.85	57.55	56.12	97.51	34781
		1000	47.76	47.68	46.60	97.73	32759
		1000	74.04	73.54	71.84	97.68	37535
		500	45.60	45.29	44.18	97.55	36510
		500	42.07	41.85	40.77	97.43	32797
		100	50.06	50.03	48.62	97.19	32916
		100	44.67	44.63	43.40	97.25	25770
	Input		24.92	24.92	24.00	96.30	

Sample type	Histone mark	No. of cells	Total reads (million)	Trimmed reads (million)	Aligned reads (million)	Alignment efficiency %	Peaks
Cere-bellum mix	H3K4me3	10000	23.59	23.41	22.63	96.69	43454
		10000	31.61	31.32	30.34	96.87	43255
		1000	24.85	24.81	23.65	95.33	46184
		1000	26.84	26.80	25.48	95.07	46419
		100	25.27	25.25	23.86	94.51	32148
		100	23.09	23.07	21.83	94.60	28570
	H3K27ac	10000	22.07	21.85	21.24	97.21	45345
		10000	27.11	26.80	26.07	97.29	48655
		1000	26.97	26.93	26.00	96.58	61435
		1000	31.93	31.85	30.75	96.55	67921
		100	30.58	30.55	29.19	95.55	47596
		100	30.74	30.71	29.36	95.61	54179
	H3K27me3	10000	47.86	47.61	45.84	96.26	8951
		10000	46.40	46.25	44.55	96.33	9792
		1000	54.51	54.40	52.10	95.77	10053
		1000	35.50	35.41	33.92	95.80	9311
		100	45.47	45.41	43.25	95.23	9457
		100	53.12	53.05	50.56	95.30	9518
	Input		28.40	28.37	27.30	96.24	
Cere-bellum mix (ENCODE)	H3K4me3	100 M	21.40	20.96	16.83	80.31	39321
			18.71	18.69	9.76	52.22	30456
	H3K27ac		13.35	13.32	12.24	91.87	17000
			11.73	11.71	10.96	93.60	34663
	H3K27me3		36.36	32.93	30.72	93.30	1357
			33.13	29.73	26.02	87.53	7806
	Input		16.35	11.85	10.60	89.44	

Sample type	Histone mark	No. of cells	Total reads (million)	Trimmed reads (million)	Aligned reads (million)	Alignment efficiency %	Peaks
Cere-bellum neuron	H3K4me3	1000	27.73	27.69	26.89	97.14	56319
		1000	23.93	23.90	23.20	97.07	55057
		100	23.54	23.52	22.43	95.36	58409
		100	21.41	21.37	20.37	95.28	59744
	H3K27ac	1000	28.25	28.21	27.58	97.78	46359
		1000	24.59	24.45	23.87	97.60	52503
		100	24.75	24.71	23.69	95.86	25059
		100	20.20	20.12	19.29	95.90	24988
	H3K27me3	1000	43.97	43.97	43.02	97.85	7215
		1000	43.16	43.15	41.94	97.21	6480
		100	43.16	43.15	41.02	95.08	4922
		100	38.07	38.06	36.18	95.05	4345
	Input		24.74	24.68	23.79	96.40	
Cere-bellum glia	H3K4me3	1000	17.30	13.43	12.32	91.79	36191
		1000	26.05	21.68	20.24	93.34	48900
		100	18.93	18.10	16.95	93.66	7345
		100	22.23	21.41	20.05	93.64	9655
	H3K27ac	1000	21.19	21.18	20.60	97.26	33849
		1000	24.17	24.13	23.47	97.28	31808
		100	22.94	22.76	21.93	96.36	6348
		100	24.94	24.86	23.98	96.48	5567
	H3K27me3	1000	34.55	34.54	33.45	96.85	11480
		1000	36.85	36.83	35.66	96.84	10532
		100	41.55	41.54	40.09	96.50	11812
		100	44.17	44.14	42.59	96.47	9539
	Input		21.99	21.96	21.19	96.47	
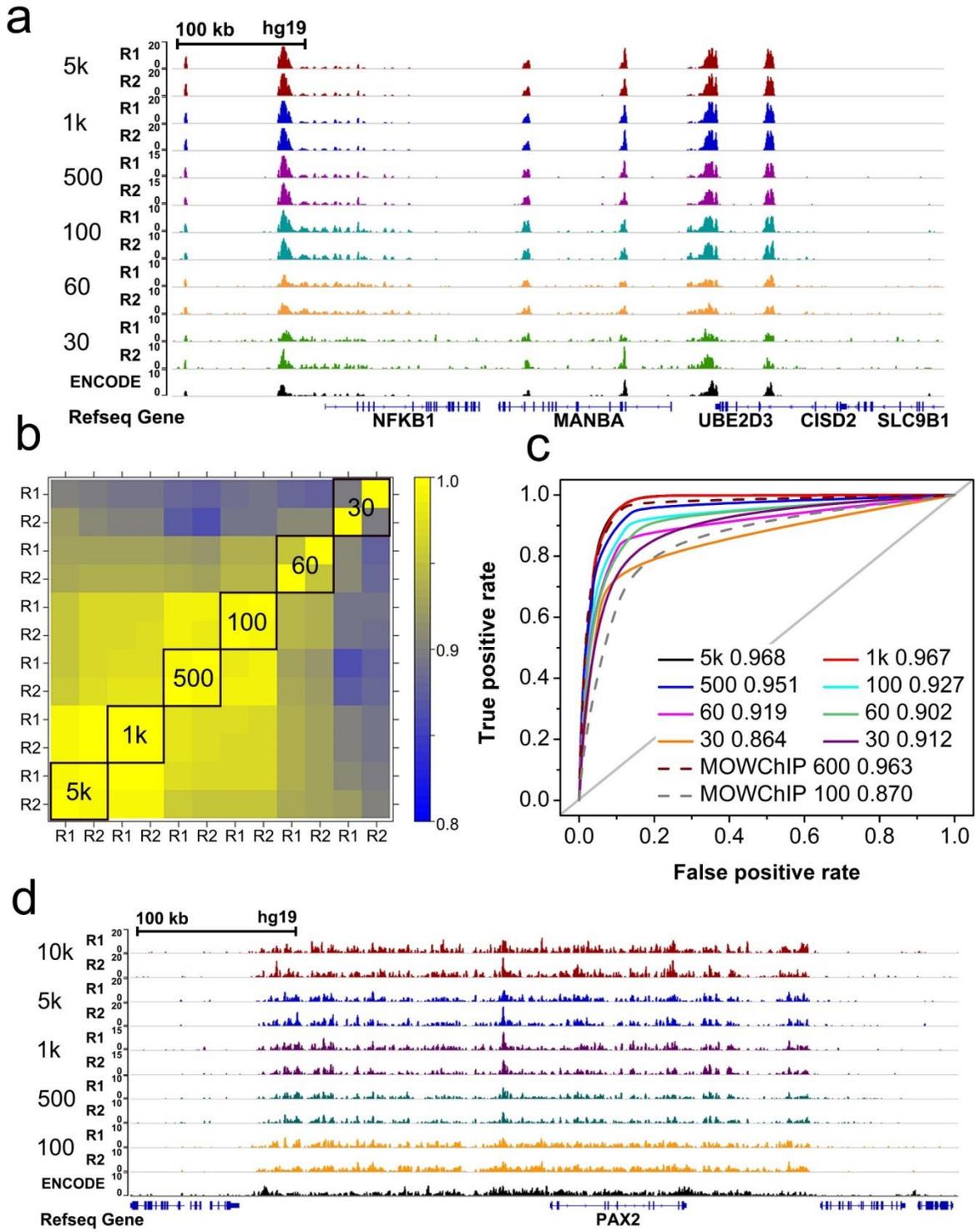
Fig. 5.6 SurfaceChIP-seq generates high quality data using as little as 30-100 cells. (a) Representative H3K4me3 signals of GM12878 using 5k to 30 cells. (b) H3K4me3 cross correlation of promoter regions. (c) ROC curves of SurfaceChIP-seq and MOWChIP-seq data for H3K4me3 mark. (d) Representative H3K27me3 signals of GM12878 using 10k to 100 cells.

Benchmark SurfaceChIP-seq.

To benchmark the performance of SurfaceChIP-Seq for genome-wide analysis, we profiled H3K4me3 and H3K27me3 marks with various numbers of GM12878 cells. We digested chromatin using 5000, 1000, 500, 100, 60 and 30 cells for H3K4me3 mark and 10000, 5000, 1000, 500 and 100 cells for H3K27me3 mark. **Fig. 5.6a** shows normalized SurfaceChIP-Seq signals for H3K4me3 at the *MANBA* gene locus. The promoter region was highly enriched for H3K4me3 signal in SurfaceChIP-Seq data generated using 30-5000 cells. We examined the reproducibility of the SurfaceChIP-seq. The H3K4me3 samples of 5000, 1000, 500, 100, 60, and 30 cells had Pearson coefficients of 0.995, 0.997, 0.992, 0.991, 0.954, and 0.897 between replicates, respectively (**Fig. 5.6b**). The Pearson coefficients significantly dropped when 60 or fewer cells were used. Using published ChIP-Seq data generated by conventional protocol with millions of cells as the gold standards, we compared the performance of SurfaceChIP-Seq to state of the art methods, MOWChIP-Seq (detailed in **Materials and Methods**). Receiver Operating Characteristic (ROC) curve was used to quantify the agreement of data with a gold standard. The Area Under the ROC Curve (AUC) is a standard metric for quantifying balanced sensitivity and specificity. For H3K4me3, SurfaceChIP-Seq with 30 cells (AUC = 0.912 and 0.864) produced data with quality that was comparable to that of

MOWChIP using 100 cells (AUC=0.870, **Fig. 5.6c**). Although the AUC values decreased with decreasing number of cells, all data showed good consistence with gold standard that enabled analysis of important genome-wide features. H3K27me3 is a repressive mark of transcription and involved in developmental transitions in both animals and plants. **Fig. 5.6d** shows normalized SurfaceChIP-Seq signals for H3K27me3 around *Pax2* gene using 100-10000 cells. *Pax2* gene provides instruction for making transcription factor Pax which is critical for formation of tissues and organs during embryonic development. The H3K27me3 samples of 5000, 1000, 500, 100, 60, and 30 cells had Pearson coefficients of 0.978, 0.971, 0.964, 0.941, and 0.960 between replicates, respectively (**Fig. 5.7**).
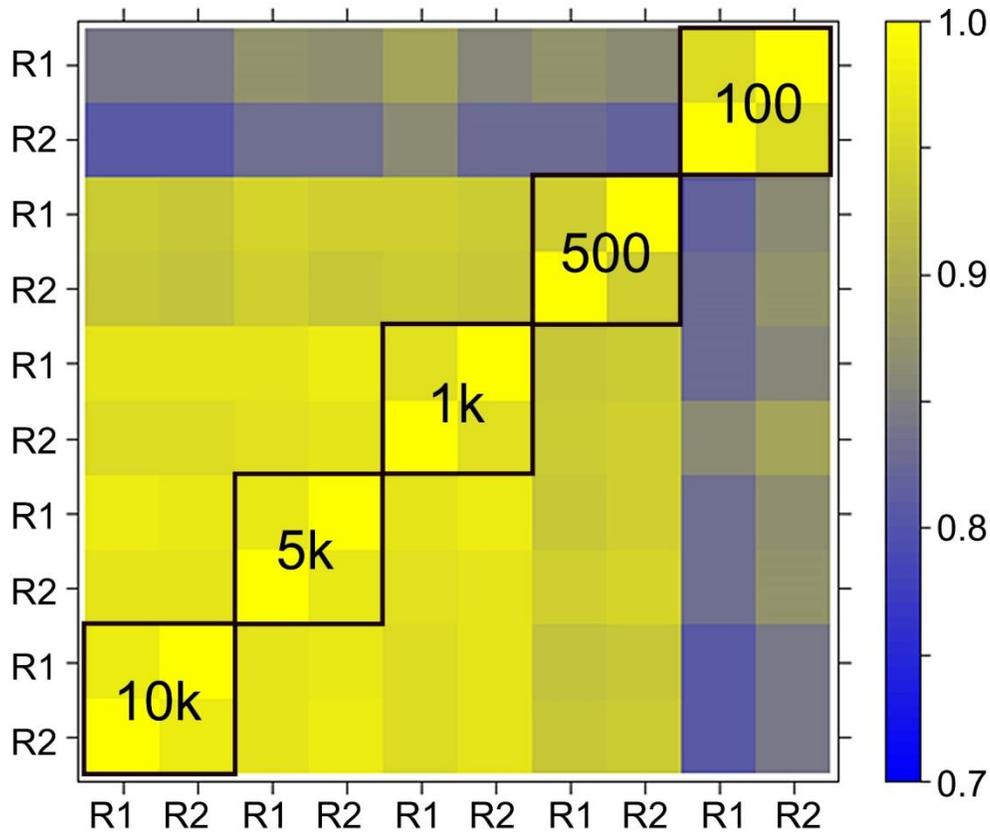
Fig. 5.7 Cross correlation of H3K27me3 signals in promoter regions from GM12878 cells.
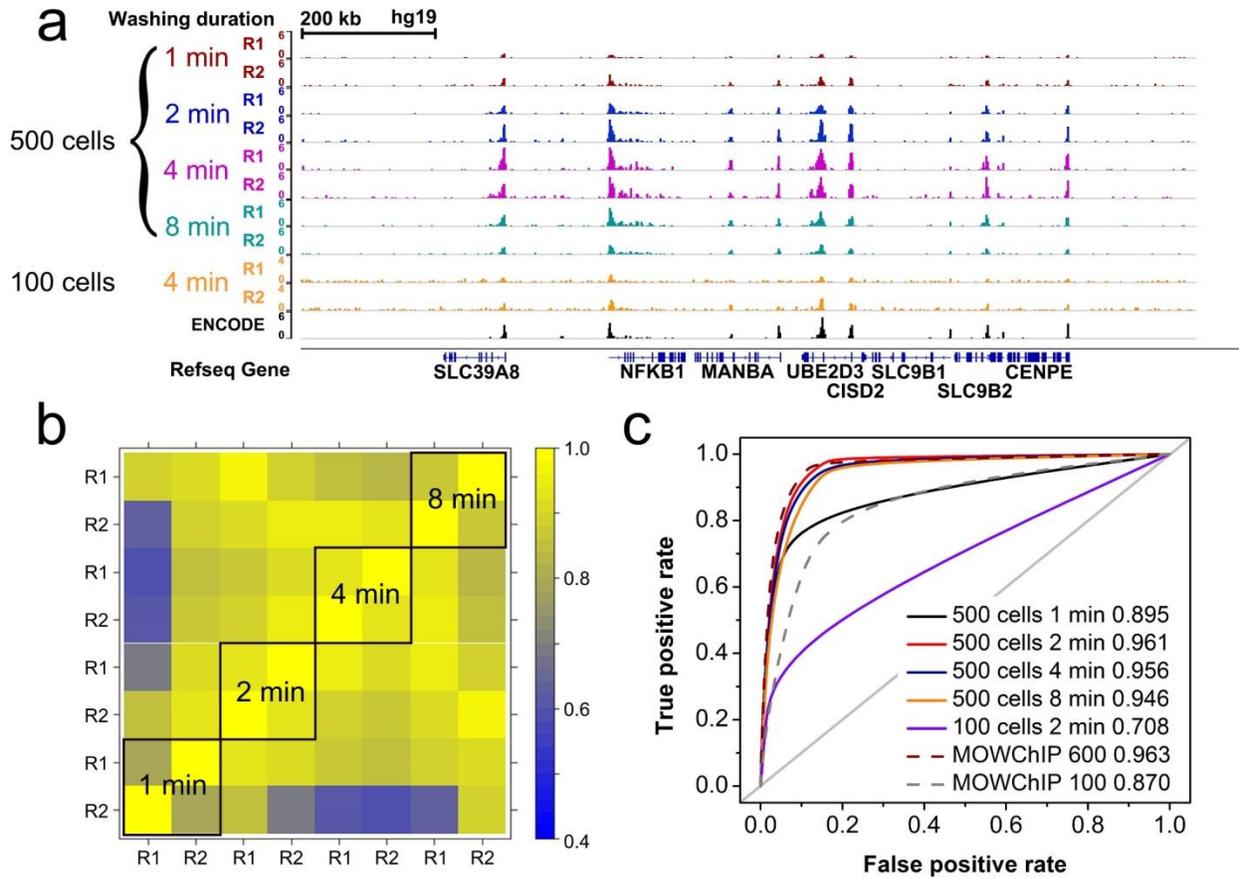
Fig. 5.8 The SurfaceChIP-seq signals using AGP linker were improved by optimizing washing duration. (a) Representative H3K4me3 signals of GM12878 using 500 or 100 cells. (b) Cross correlation of promoter regions. (c) ROC curves of SurfaceChIP-seq and MOWChIP-seq data for H3K4me3 mark.

As an alternative to PLL/oligo linker, we tested (3-Aminopropyl) triethoxysilane /Glutaraldehyde/Protein A linker (AGP) linker using chromatin digested from 500 cells for H3K4me3 mark (**Fig. 5.8a**). The oscillatory washing duration was optimized (1, 2, 4 and 8 min) to effectively remove nonspecific adsorption without compromising peak quality. 2 min and 4 min washing show the highest Pearson coefficients (0.922 and 0.926 between replicates, respectively, **Fig. 5.8b**) and AUC (0.961 and 0.956, respectively, **Fig. 5.8c**). This is consistent with the qPCR optimization of washing duration (**Fig. 5.4**). SurfaceChIP-Seq with AGP linker using 500 cells produced data with quality that was comparable to that of MOWChIP-seq using 600 cells (AUC=0.963, **Fig. 5.6c**). Using 100 cells, the AGP linker (AUC=0.708) underperformed PLL/oligo linker (AUC=0.927) and MOWChIP-seq (AUC=0.870) possibly due to the less antibody coating density.

Histone modification landscapes are distinctive for neurons and glia cells.

We applied SurfaceChIP-Seq to study the epigenome of neurons (NeuN+) and glial cells (NeuN-) isolated from mouse cerebellum. Little is known about the epigenomic signature of the subpopulations from cerebellum, largely due to the difficulty in isolating sufficient qualities of intact chromatin after harsh ultracentrifugation and FACS sorting. Our technology provides an ideal solution to this challenge. Neuron and glial cells are the two major populations in the brain with distinctive functional and epigenomic patterns. Although, previous studies have mapped histone modifications to the entire cerebellum [194] and subpopulations from prefrontal cortex [195], the neurons and glial cells from cerebellum have not been investigated. We mapped H3K4me3, H3K27ac, and H3K27me3 using 10000, 1000, and 100 nuclei from cerebellum mixture and 1000 and 100 nuclei from neuron and glial cells (**Table 5.2**). For all histone marks and sample sizes,

the replicates are highly correlated (average r= 0.97, 0.96, and 0.97 for H3K4me3, H3K27ac, and H3K27me3, **Fig. 5.9**). Lower correlations were observed among two different populations indicates the different histone modification patterns (average r= 0.88, 0.86, and 0.81 for H3K4me3, H3K27ac, and H3K27me3, **Fig. 5.9 and 5.10**). For example, the Pearson correlation between H3K4me3 mark in neuron replicates using 100 nuclei is 0.97 which is much higher than the correlation between neuron and glial cells (r = 0.86, **Fig. 5.11a and 5.11b**). We also computed the correlation between replicates of our data with published ENCODE datasets (~100 million cells) on entire cerebellum. Our data showed similar or better correlation between replicates compared to published data which supports the high quality of our data. (average r = 0.98 and 0.91 for our data and ENCODE datasets, respectively).
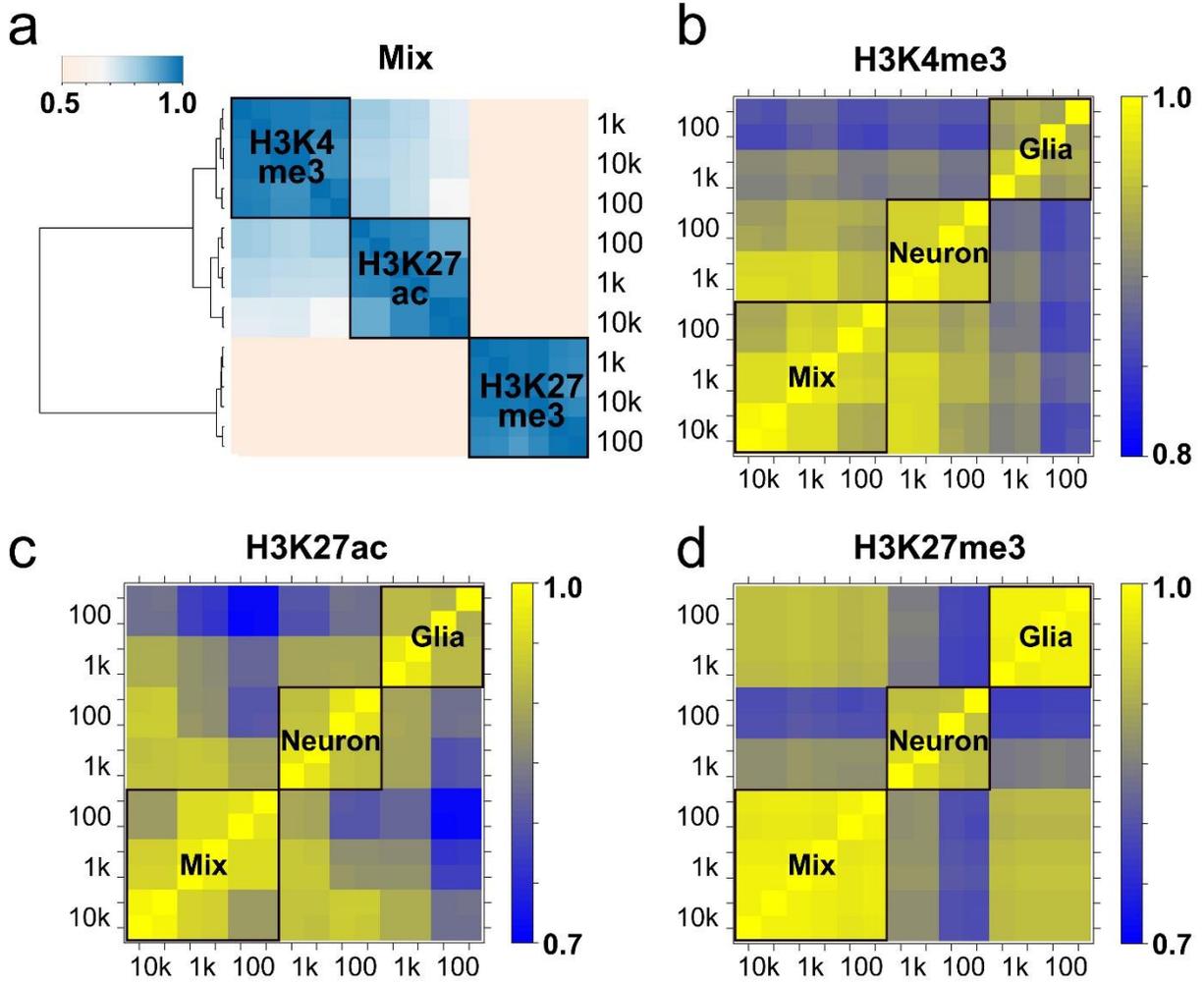
Fig. 5.9 The cross correlation of SurfaceChIP-seq signals from mouse samples. (a) Clustering of ChIP-seq signals from mix using various number of cells. (b-d) Cross correlation of promoter coverage from various cell types.
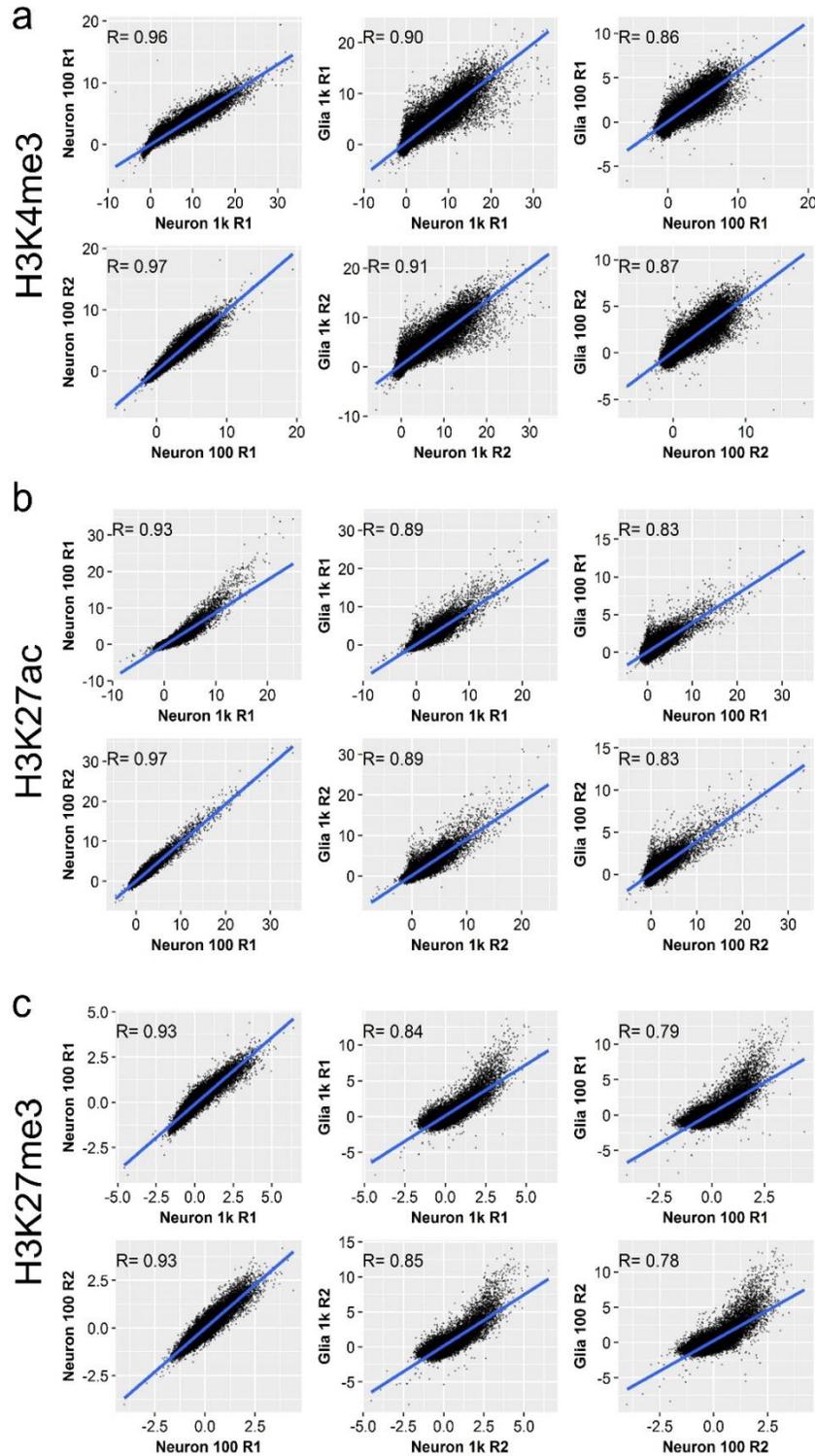
Fig. 5.10 The scatter plots of SurfaceChIP-seq signals from mouse samples. (a-

c) show the plots for H3K4me3, H3K27ac and H3K27me3 marks, respectively.
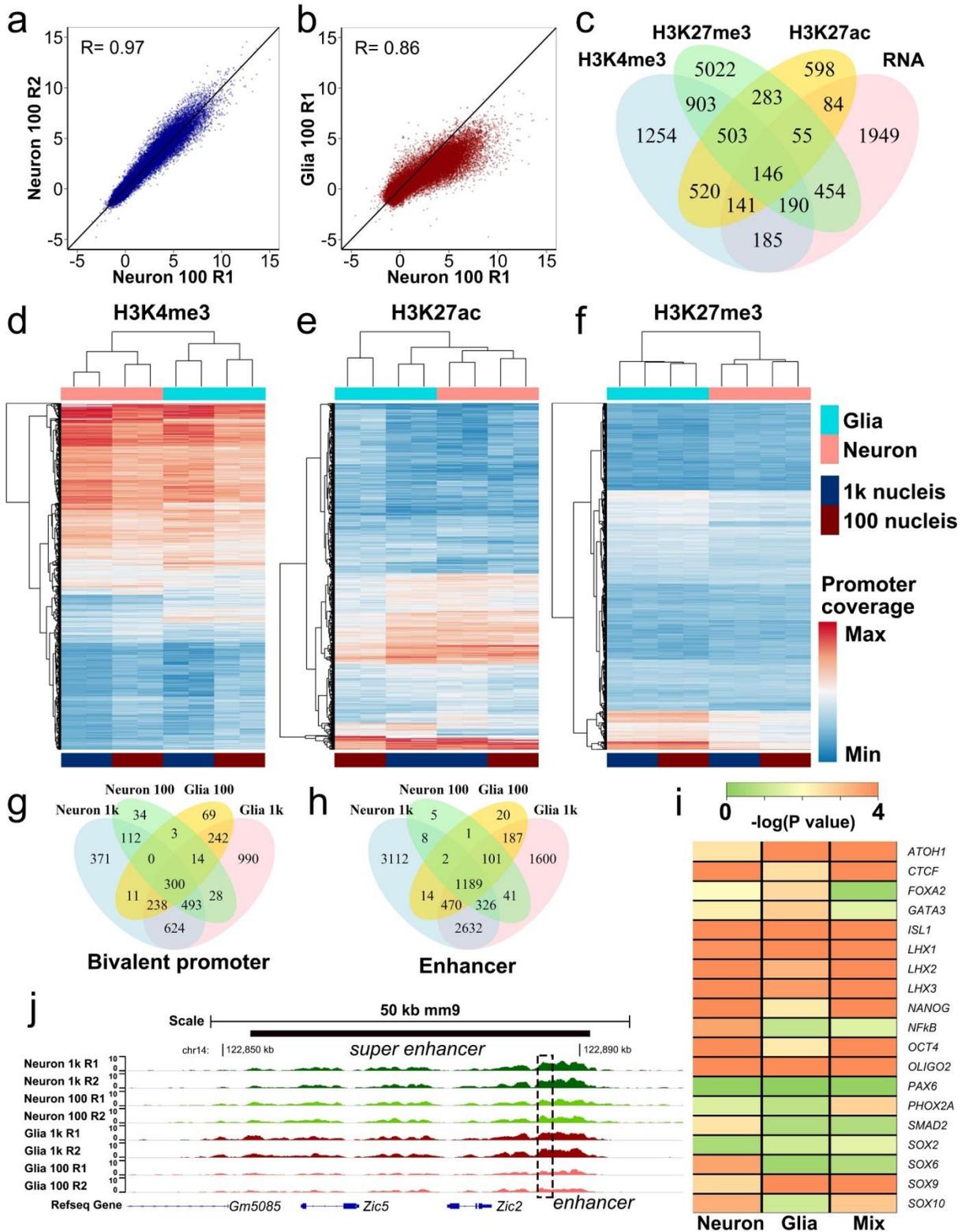
Fig. 5.11 The distinctive histone modification profile between Neuron and Glia.

(a) H3K4me3 promoter signals on neuron replicates using 100 nuclei. (b) H3K4me3 promoter signals on neuron and glia using 100 nuclei. c) Venn diagram of differentially binding promoters (H3K4me3, H3K27me3 and H3K27ac, FDR < 0.05) and differently expressed genes (RNA, p value < 0.01, fold change > 8) (d-f) Unsupervised hierarchical clustering of genome-wide coverage for the significantly binding H3K4me3. (g) Venn diagram of bivalent promoters from each cell type. (h) Venn diagram of enhancers from each cell type. (i) Heat map of p values for representative transcription-factor-binding motifs that are enriched in enhancers (j) Representative super enhancer around Zic2 and Zic5 loci which is related to neural crest differentiation and transcription factor activity.

Using our technology, we distinguished the histone modification patterns between the two populations even with such little starting materials. For example, H3K4me3 mark was enriched at *RTNK* gene in glial cell but not in neurons (**Fig. 5.11c**). We identified 3842, 2330, and 7556 promoter regions with significantly different enrichments (FDR < 0.05) from H3K4me3, H3K27ac, and H3K27me3 marks, respectively (**Fig. 5.11c**). Among them, 2092 promoters showed differences on two histone marks and 649 promoters showed differences on all three histone marks. We also performed RNA-seq from nuclei to validate our results (**Chapter 4**). The differentially enriched promoters largely overlapped with genes showed significantly different expression levels between two populations (1255 out of 3204 genes, p value < 0.01, fold change >8). We observed positive correlations (H3K4me3 and H3K27ac, **Fig. 5.12 and 5.13**) and negative

correlation (H3K27me3, **Fig. 5.14**) between promoter coverages and gene expression levels. The result supports the report that H3K4me3 mark is associated with activating transcription, while H3K27me3 is a repressive mark. By comparing the coverages of promoters with significantly different enrichment, dramatically different patterns were observed (**Fig. 5.11e-g**). Unsupervised hierarchical clustering of H3K4me3 and H3K27me3 signals successfully identified the two populations.
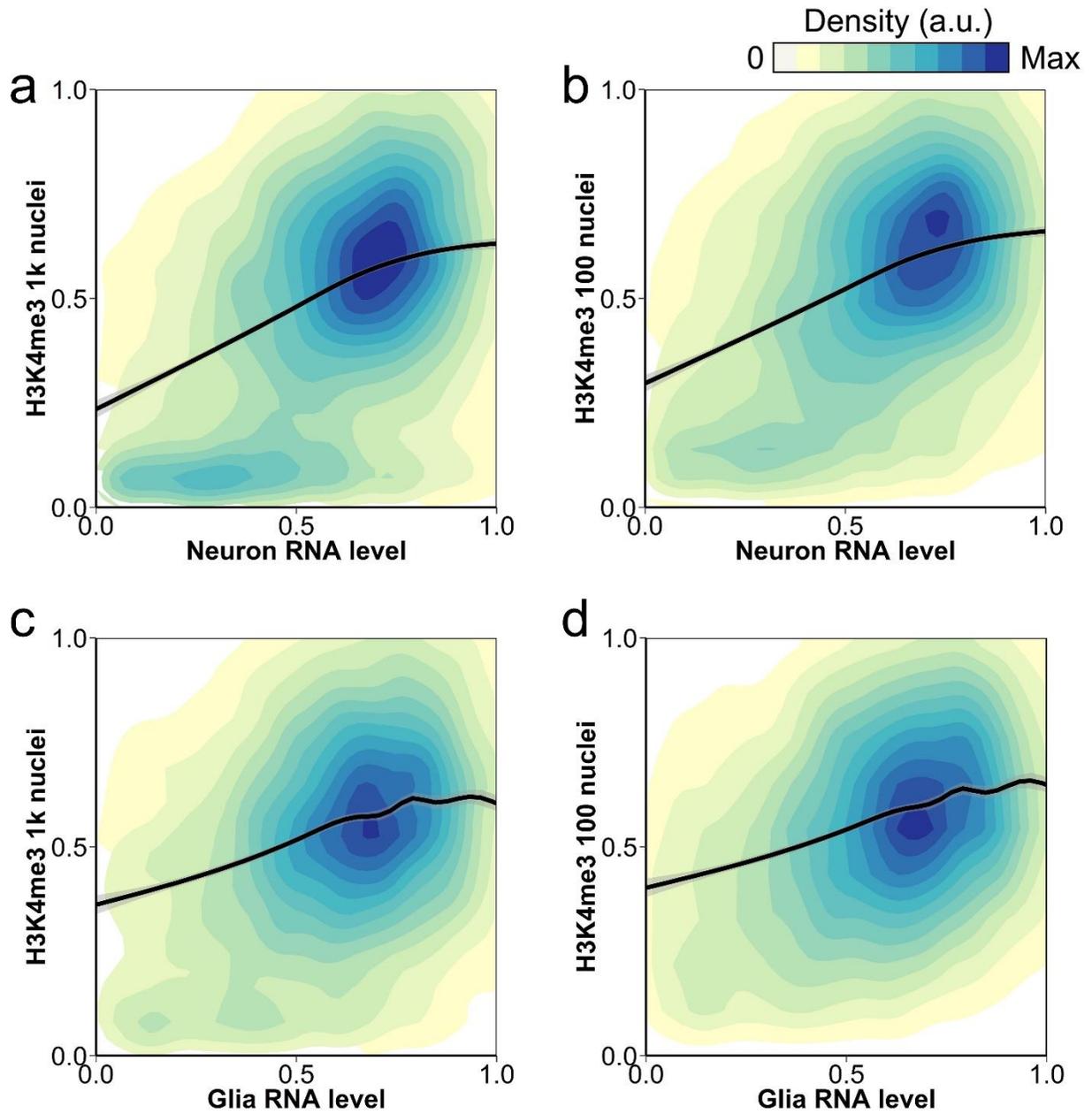
Fig. 5.12 The positive correlation between promoter signal of H3K4me3 mark and mRNA expression level.
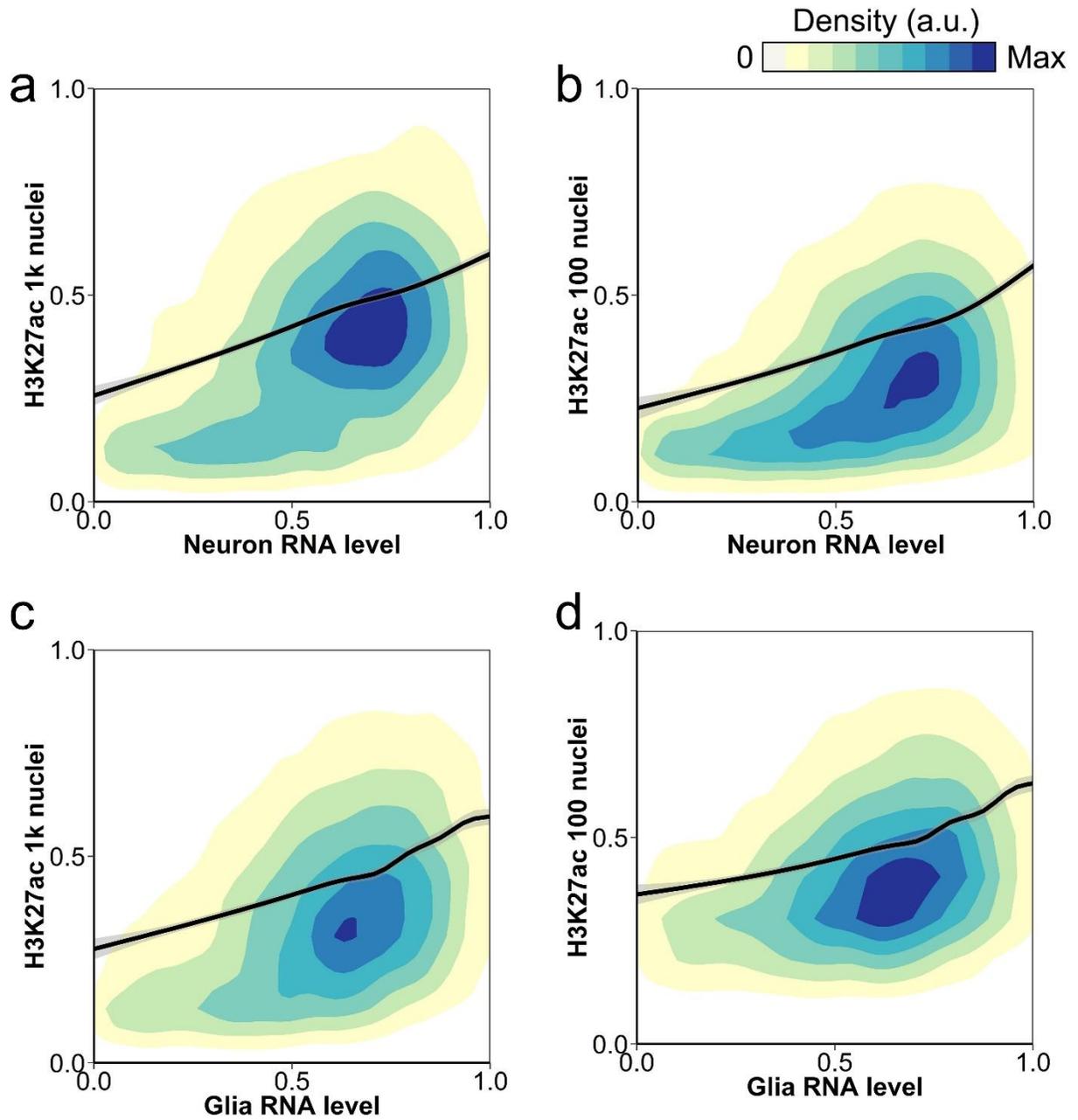
Fig. 5.13 The positive correlation between promoter signal of H3K27ac mark and mRNA expression level.
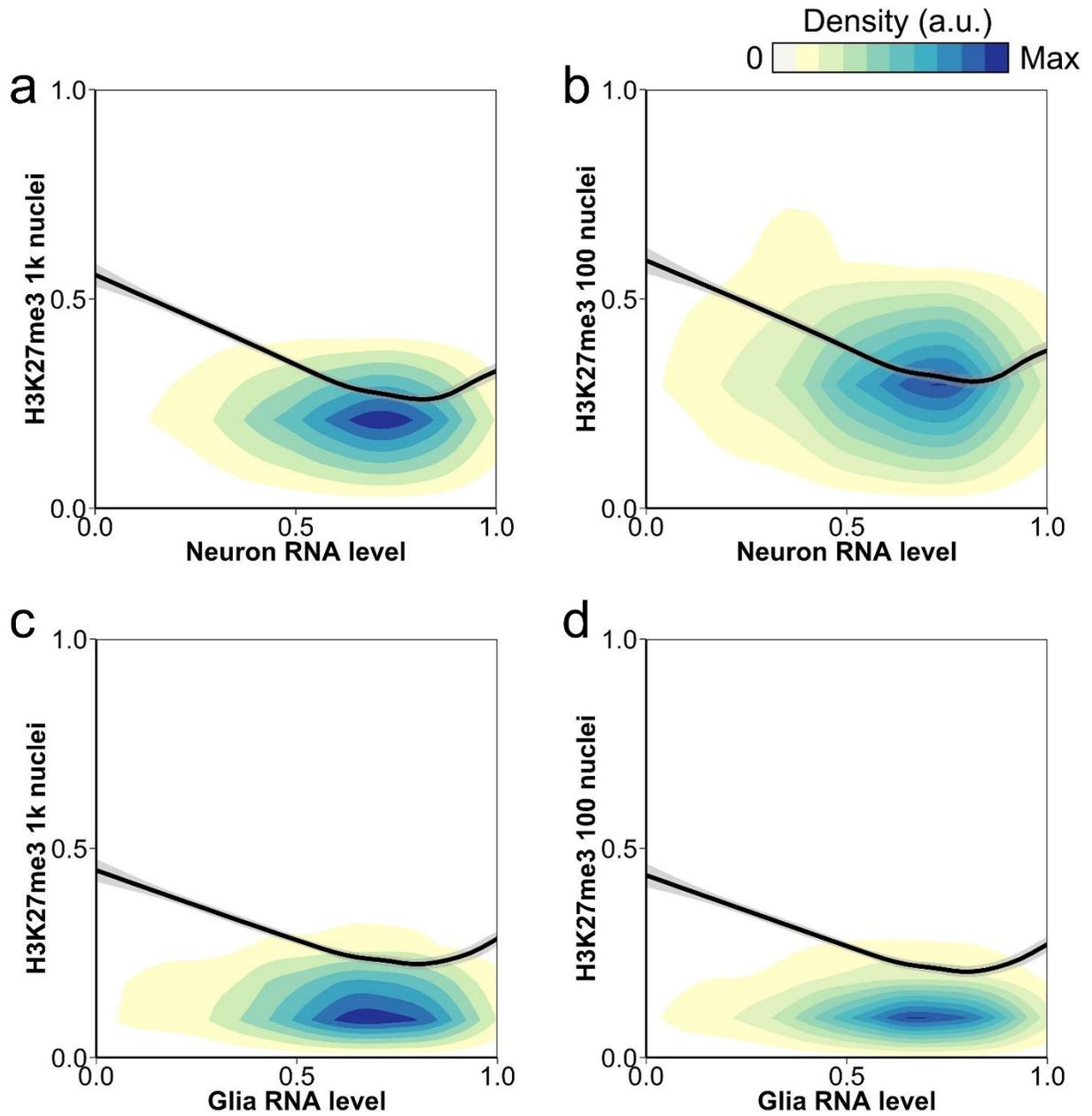
Fig. 5.14 The negative correlation between promoter signal of H3K27me3 mark and mRNA expression level.

Bivalent promoter is a distinctive histone modification signature that combines activating H3K4me3 mark and repressive H3K27me3 mark. It is believed to play critical roles in pluripotency by poising the expression of developmental genes [196]. Using SurfaceChIP-seq data with 1000 nuclei, we discovered 2837, 2149, and 2929 bivalent promoters from mix, neuron, and glia, respectively (**Fig. 5.11h**). Using data generated by 100 nuclei, we discovered 2534, 984, and 877 bivalent promoters from mix, neuron, and glia, respectively. Averagely 89.2 ± 2.9% of the 100-nuclei bivalent promoters are overlapped with 1000-nuclei bivalent promoters which showed high consistence. We noticed that the bivalent chromatin masks developmental regulatory genes in both cell types. To compare the cell-type specific bivalent promoters, we performed gene ontology (GO) term enrichment analysis. Interestingly, we found that the Homeobox-related functions are highly enriched in neuron-specific bivalent promoters (P values of Homeobox = 3.3×10^{-10} and 6.8×10^{-11} for 1000 and 100 nuclei datasets, respectively), but not in glia-specific bivalent promoters. Homeobox genes are known to be critical in neuronal development [197].

Enhancers are short regions that can be bound by activators to increase the transcription of a particular gene. Little is known about the enhancers in the cerebellum. Using SurfaceChIP-seq data, we predicted the active enhancers in both neuron and glial cells. We predicted 7753, 1673, 6546, and 1984 enhancers using 1000-neuron-nuclei, 100-neuron-nuclei, 1000-glia-nuclei, and 100-glia-nuclei datasets, respectively (**Fig. 5.11i)**. There are 1189 enhancers shared among the four datasets. 3125 and 1797 enhancers are specific to neuron and glial cells, respectively. Using 100 nuclei datasets, we identified averagely 25% of the enhancers. The DNA motif search revealed that the

enhancers in neuron and glial cells are not identically enriched for known transcription factors (**Fig. 5.11j**). For example, *CTCF*, *OCT4*, and *SOX6* are specifically enriched in neuron, while *ATOH1* and *SOX9* are specifically enriched in glial cells.

Super enhancer (SE) is a class of regulatory regions with unusually strong enrichment for the transcriptional coactivator binding. They are typically much longer than single enhancers and play critical roles in cell differentiation and neuron function. There is very little knowledge about the SE in entire cerebellum, not to mention the neuronal and glial subpopulation. Using our SurfaceChIP-seq data, we identified 602 and 314 SEs that are specific to neuron and glia, respectively (**Fig. 5.15**). The target genes of our SEs are enriched in neuronal function, such as synapse (p value < 10^{-11}, hypergeometric test), neurogenesis (p value < 2×10^{-6}), and neuron migration (p value < 2×10^{-5}). These target genes include many known neuronal function regulators, such as *Rims1*, *Cadm3*, *Paqr8*, *Zic2* and *Zic5* [198] (**Fig. 5.11k**).
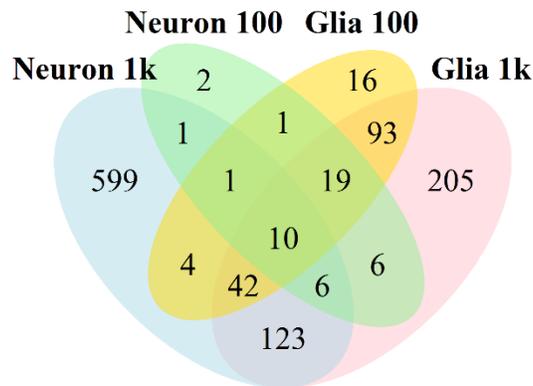
Fig. 5.15 Venn of super enhancer overlap between various samples.

5.3 Materials and Methods

Device fabrication

1) Single-layer PDMS device. The polydimethylsiloxane (PDMS)-glass chip was fabricated by soft lithography [29, 146, 177]. The photomask with the microscale patterns was designed using LayoutEditor (juspertor GmbH) and printed on high-resolution (10,000 d.p.i.) transparencies. The pattern on photomask was replicated on a 3-inch silicon wafer (P(100), 380 μm thickness, University Wafers) by spinning SU-8 2025 (Microchem) on at 500 rpm for 10 s followed by at 1500 rpm for 30 s (yielding 60 μm in SU-8 2025 thickness). The PDMS (RTV615, R S Hughes) with a mass ratio of A: B= 10: 1 was poured onto a master in petri dish (~4 mm thickness). The PDMS was baked at 75 ℃ for 1 h to cure. The cured PDMS was peeled off from the master and access holes were punched.

2) Multi-layer PDMS device. Three photomasks (1 for the control layer and 2 for the fluidic layer) were printed on high-resolution (10,000 d.p.i.) transparencies. The control layer master was fabricated by spinning SU-8 2025 on a silicon wafer at 500 rpm for 10 s followed by at 3000 rpm for 45 s (yielding 24 μm in SU-8 2025 thickness). The fluidic layer master was fabricated in SU-8 2025 and AZ 9260 (Clariant) on another silicon wafer with the thickness being 60 and 21 μm, respectively. The fluidic layer master was baked at 130 ℃ for 30 s to round AZ 9260 features so that the resulted fluidic channels could be fully closed by microscale pneumatic valves. The control layer (~0.4 cm thick) was fabricated by pouring PDMS prepolymer with a mass ratio of A: B =5: 1 onto the control layer master in a petri dish. The fluidic layer (~80 μm thick) was fabricated by spinning PDMS (A: B=20: 1) at 500 rpm for 10 s and then at 1500 rpm for 30 s. Both layers were

baked at 75 ℃ for 15 min. The control layer was then peeled off from the master and access holes to the control layer channels were punched. The two PDMS layers were aligned, brought into contact and baked for 1 h at 75℃ for thermal boning. The PDMS structure was peeled off from the fluidic layer master and access holes to the fluidic layer were punched.

The PDMS structure (single or multi-layer) was bonded to a glass slide that was pre-cleaned in 27% NH_4OH: 30% H_2O_2: H_2O=1: 1: 5 (volumetric ratio) at 80 ℃ for 30 min. Both PDMS surface and glass surface were oxidized in plasma (PDC-32G, Harrick Plasma) for 1 min and then immediately brought into contact. The bonded device was immediately used for antibody coating.

Antibody-oligo conjugate synthesis. Antibodies and DNA oligo (5'-NH_3-AAA AAA AAA ATA CGG ACT TAG CTC CAG GAT-3') was conjugated using Protein-Oligo Conjugation Kit (Solulink) following kit instruction. The reagent volumes were determined for each reaction based on oligo/antibody concentration and amount using a spreadsheet calculator provided by the kit. Antibody and DNA oligo were conjugated with S-HyNic (succinimidyl-6-hydrazino-nicotinamide) linker and S-4FB (succinimidyl-4-formylbenzamide) linker, respectively. DNA oligo was suspended to 0.5 OD_{260}/µl by oligo resuspension solution and quantified by Evolution 300 UV-Vis spectrophotometer (Thermo Fisher) with micro UV cuvettes (Z637092, Sigma-Aldrich). 30-60 µl oligo was buffer exchanged to 1× modification buffer using Zeba column (Solulink). 1 mg S-4FB Reagent was dissolved with 40 µl anhydrous DMF. DMF, dissolved S-4FB and oligo was

mixed and incubated at room temperature for 2 h for oligo modification. The modified

oligo was buffer exchanged to 1× conjugation buffer using Zeba column. The

concentration of 4FB-Oligo was quantified by UV-Vis spectrophotometer. The molar

substitution ratio (i.e. number of linker per target molecules) was checked using 2-HP

reagent to ensure successful modification. The antibody was concentrated to 50-130 μl

with concentration of 1-5 mg/ml and buffer exchanged to 1× modification buffer. The

antibody was incubated with S-HyNic reagent at room temperature for 2.5 h. The HyNic-

antibody molar substitution ratio was checked using 2-SBA reagent. HyNic-antibody and

4FB-Oligo was mixed and incubated at room temperature for 3 h for conjugation. The

antibody-oligo conjugate was purified by Zeba column and buffer changed to 1× PBS at

stored at 4 ℃. The concentration of antibody-oligo was quantified by Pierce BCA Protein

Assay Kit (Thermo Fisher).

Antibody coating. The glass slide was pre-cleaned in 27% NH$_4$OH: 30% H$_2$O$_2$: H$_2$O=1: 1:

5 (volumetric ratio) at 80 ℃ for 30 min. The glass was blown dried by air. Both PDMS

surface and glass surface were immediately oxidized in plasma (PDC-32G, Harrick

Plasma) for 1 min and then immediately brought into contact. The bonded chip was

immediately used for either following protocol.

1) **PLL/oligo linker**

The microfluidic channels and reservoirs were filled with 0.1% poly-l-lysine (P2636,

Sigma-Aldrich) and sealed with microseal 'B' adhesive seals (MSB1001, Bio-Rad). The

chip was incubated at 37 ℃ for 15 min. After incubation, the channels were washed by

100 µl H$_2$O and incubated at 60 ℃ for 60 min. The microfluidic channel was then filled

with 200 µM coating DNA (5'-AAA AAA AAA AAA AAT CCT GGA GCT AAG TCC GTA-

3') dissolved in DMSO: H$_2$O=1: 2 (volumetric ratio). The microfluidic chip was then kept

in desiccator at room temperature for about 3 days until all the solvent evaporated. The

chip was then baked at 75 ℃ for 3 h to promote crosslinking. The microfluidic channel

was washed with 100 µl blocking buffer to reduce non-specific adsorption. The channel

was incubated with blocking buffer for 60 min at 37 ℃. Antibody-oligo conjugate was

diluted in blocking buffer (3% BSA, 0.05% Tween 20 in PBS buffer) to desired

concentration and then loaded in the channel. For H3K4me3 and H3K27ac marks, 50

µg/ml, 25 µg/ml and 10 µg/ml antibody-oligo conjugate was used for 5000, 1000-500 and

less than 500 cells, respectively. For H3K27me3, 100 µg/ml, 50 µg/ml and 25 µg/ml

antibody-oligo conjugate was used for 10k-5k, 1k-500 and 100 cells, respectively. The

chip was incubated at 37 ℃ for 60 min.

To examine the coating density, 50 µM FAM labelled DNA oligo (5'-NH$_3$-AAA AAA AAA

ATA CGG ACT TAG CTC CAG GAT-3') dissolved in blocking buffer was incubated in

oligo coated chip for 60 min at 37 ℃ and washed with 100 µl blocking buffer. The

fluorescence was checked to ensure uniform coating.

2) AGP linker

Freshly prepared 10% 3-aminopropyltriethoxysilane (APTES, 440140, Sigma-

Aldrich)/ethanol was injected into microfluidic channels. The reservoirs were fully filled

with 10% APTES/ethanol (~25 µl each channel) and sealed with microseal 'B' adhesive

seals to reduce contact with moisture in the air. The chip was incubated at room temperature for 30 min for surface silanization. Each channel was washed with 100 µl freshly prepared 95% ethanol. The ethanol was dried by air at 10 psi controlled by pneumatic valves. The dried chip was baked at 125 ℃ for 30 min to remove residual ethanol and promote silanization. The microfluidic channels and reservoirs are filled with 8% glutaraldehyde (G7651, Sigma-Aldrich) and sealed by microseal 'B' adhesive seals. The chip was incubated at room temperature for 30 min. The channels were washed by 100 µl water and dried by air at 10 psi. The channel was coated with 2 mg/ml ProteinA/PBS (21184, Life Technologies) for 60 min and washed by 100 µl 0.5% BSA/PBS. 10 µl of 200 µg/ml antibody diluted by 0.5% BSA/PBS was loaded in each channel. The reservoirs were sealed by microseal 'B' adhesive seals. The chip was incubated at 4 ℃ overnight.

After antibody coating by either protocol, the channels were washed by 100 µl of 1× lysis buffer (2% Triton X, 50 mM Tris pH 7.5, 50 mM NaCl and 15 mM $MgCl_2$) with freshly added 1 µl protease inhibitor cocktail (PIC, P8340, Sigma-Aldrich) and 1 µl of 100 mM PMSF. The channel was washed again with 100 µl 0.5% BSA/PBS and ready for immunoprecipitation. Extra care was taken to avoid introducing bubbles into channels.

Cell culture. GM12878 cells were purchased from Coriel Institute for Medical Research. The cells were cultured in RPMI 1640 medium (11875-093, Gibco) with 15% fetal bovine serum (26140-079, Gibco) and 100 U/ml penicillin-streptomycin (15140-122, Gibco) at

37 ℃ in a humidified incubator with 5% CO_2. Cells were maintained in exponential growth phase by sub-culturing every 2-3 days.

Mouse strain and cerebellum dissection. The mouse line C57Bl/6 was bred and maintained in groups of 3-4 per cage with food ad libitum. Salk Institute's Institutional Animal Care and Use Committee (IACUC) approved all the protocols. 8-10 week old mice were anesthetized with isoflurane and decapitated. The cerebellum was rapidly dissected and frozen on dry ice. The cerebellum was stored at -80 ℃ until used for nuclei isolation.

Nuclei isolation from mouse cerebellum. All the procedure was performed on ice. All centrifugation was performed at 4 ℃. One cerebellum was placed in 5 ml ice-cold nuclei extraction buffer (0.32 M sucrose, 5 mM $CaCl_2$, 3 mM $Mg(Ac)_2$, 0.1 mM EDTA, 10 mM Tris-HCl and 0.1% Triton X-100) with freshly added 50 µl protease inhibitor cocktail (P8340, Sigma-Aldrich), 5 µl 100 mM PMSF and 5 µl 1 M DTT. 7.5 µl 40 U/µl RNase inhibitor (N2611, Promega) was mixed with nuclei extraction buffer when RNA-seq was conducted. Once thawed, the tissue was homogenized by slowly doucing for 15 times with a loose pestle and 25 times with a tight pestle (D9063, Sigma-Aldrich). The homogenate was filtered with 40 µm nylon mesh cell strainer (22363547, Fisher Scientific) and transferred into a 15 ml centrifugation tube. The sample was centrifuged at 1000 g for 10 min. The supernatant was removed and the cell pellet was gently resuspended in 1 ml cold nuclei extraction buffer with freshly added 10 µl protease inhibitor cocktail, 1 µl 100 mM PMSF and 1 µl 1 M DTT. 1.5 µl 40 U/µl RNase inhibitor was mixed with nuclei extraction buffer when RNA-seq was conducted. The 1 ml sample suspension was split

into two 1.5-ml micro-centrifuge tubes (500 µl each tube). Each 500 µl sample was mixed with 0.75 ml of 50% iodixanol to yield 30% iodixanol solution. The 50% iodixanol was prepared by adding 0.8 ml diluent (150 mM KCl, 30 mM MgCl$_2$, and 120 mM Tris-HCl, pH 7.8) to 4 ml 60% iodixanol (D1556, Sigma). The 30% iodixanol solution was centrifuged at 10,000 g for 20 min. After centrifugation, the supernatant was completely removed. The nuclei in each tube were incubated on ice for 10 min in 0.5 ml 2% normal goat serum (50062Z, Life Technologies) diluted by Dulbecco's phosphate-buffered saline (14190144, Life Technologies). The nuclei were resuspended and pooled (~1 ml in total). 1.5 µl 40 U/µl RNase inhibitor was added when RNA-seq was conducted. 16 µl 2 ng/µl anti-NeuN antibody conjugated with Alexa 488 (MAB377X, EMD Millipore) was incubated with nuclei for at least 1 h at 4 ℃ on an end-to-end rotator. 50k to 100k nuclei were used as non-staining control without adding anti-NeuN antibody. The stained samples were sorted using a FACS (BD FACSAria, BD Biosciences).

Chromatin digestion. 0.1 µl PIC and 0.1 µl of 100 mM PMSF was freshly added in 10 µl cell suspension. The sample was mixed with 10 µl 2× lysis buffer (4% Triton X, 100 mM Tris pH 7.5, 100 mM NaCl and 30 mM MgCl$_2$) and incubated at room temperature for 10 min. 1 µl of 0.1 M CaCl$_2$ and 2.5 µl of 10 U/µl MNase (88216, ThermoFisher) were rapidly mixed with sample and incubated at room temperature for 10 min. 2.22 µl of 0.5 M EDTA (pH 8) was mixed with sample and incubated on ice for 10 min. The sample was centrifuged at 16,100 g at 4 ℃ for 5 min. The supernatant (~24 µl) was transferred to a new micro-centrifuge tube and stored on ice.

Operation of the microfluidic device. The on-chip pneumatic valves (25 psi) and fluidic valves (5 psi) were actuated by solenoid valves (18801003-12V, ASCO Scientific) that were connected to a compressed air outlet and operated by a LabVIEW program via a computer and a data acquisition card (PCI-6509, National Instruments). Both pneumatic valve and fluidic valve were connected to microfluidic channel via a long perfluoroalkoxyalkane (PFA) high purity tubing (<u>1622L, length: 1.5 m, ID: 0.02 in. and OD: 0.0625 in.,</u> IDEX Health & Science), a Clear C-Flex tubing (EW-06422-01, ID: 0.0313 in. and OD: 0.0938 in.) and a short PFA tubing (length: 1 cm).

1) Multi-layer device

The control layer channels were filled with DI water before experiments. The on-chip pneumatic valves were actuated at 25 psi by solenoid valves (18801003-12V, ASCO Scientific) that were connected to a compressed air outlet and operated by a LabVIEW program via a computer and a data acquisition card (PCI-6509, National Instruments). The solenoid valves were connected to microfluidic channels via perfluoroalkoxyalkane (PFA) high purity tubing (<u>1622L, ID: 0.02 in. and OD: 0.0625 in.,</u> IDEX Health & Science) and a Clear C-Flex tubing (EW-06422-01, ID: 0.0313 in. and OD: 0.0938 in.). Coating reagents were delivered into the microfluidic device via PFA tubing with flow driven by a syringe pump. Chromatins and reagents for immunoprecipitation were delivered via tubing connected with solenoid valves.

Freshly digested chromatin (24 µl) was periodically flowed through microfluidic chamber (2.4 µl). The chromatin solution was driven by solenoid valve (5 psi) every 6 min to refresh the sample in the chamber. The entire sample was flowed through within 10 6-min periods. After ChIP, 20 µl of low-salt washing buffer (10 mM Tris-HCl pH 7.4, 50 mM NaCl, and 1 mM EDTA pH 8.0 with freshly added 1 mM PMSF, and 1% PIC) was flowed through microfluidic chamber. Oscillatory washing was conducted (4 min unless otherwise noted) to remove non-specifically adsorbed or physically trapped debris. We filled each Clear C-Flex tubing (connected to the chip reservoir) with 20 µl of low-salt washing buffer. Pressure pulses (at 5 psi with 1 s duration) were applied alternatingly at either end of the microfluidic chamber. The pulse duration was controlled by LabVIEW program and two solenoid valves for each chamber. We performed another round of washing with high-salt washing buffer (10 mM Tris-HCl pH 7.4, 100 mM NaCl, and 1 mM EDTA pH 8.0 with freshly added 1 mM PMSF, and 1% PIC) with the same protocol.

After washing, 10 mg/ml proteinase K was freshly diluted to 1 mg/ml by elution buffer (10 mM Tris-HCl pH 7.4, 50 mM NaCl, 10 mM EDTA pH 8.0, and 0.03% SDS). 25 µl of 2 mg/ml proteinase K was flowed through microfluidic chamber. The on-chip pneumatic valves were actuated and the chambers were isolated. The reservoirs were sealed by microseal 'B' adhesive seals to avoid cross contamination and water evaporation. The chip was placed on a hot plate and incubated at 60 ℃ for 30 min to allow the digestion of histone and antibody. After proteinase K digestion, the sample was eluted by flushing the chamber with 100 µl water.

2) Single-layer device

During antibody coating and elution procedure, a 200 µl pipette tip was attached to inlet reservoir and used to store reagents. The reagents flew through the microfluidic channel driven by a 200 µl pipette attached to the outlet reservoir. During the other steps (immunoprecipitation and washing), the single-layer device was operated the same as multi-layer device.

Extraction of ChIPed DNA. 100 µl of elute was purified by phenol extraction and ethanol precipitation. The DNA pellet was resuspended in 20 µl EB buffer. When PLL/oligo linker was used, the DNA was further purified by 32 µl SPRIselect beads and resuspended in 20 µl EB buffer. The DNA was stored at -80 ℃ until use. DNA concentrations were measured using a Nanodrop 3300 with Quant-iT PicoGreen dsDNA Assay Kit (P11496, Life Technologies).

Extraction of input DNA. Input samples (10 µl, 1000 cells) were mixed with 4 µl of 10 mg/ml proteinase K and 90 µl elution buffer and incubated at 65 ℃ for 1 h. The input sample was purified by phenol extraction and ethanol precipitation. The DNA pellet was resuspended in 20 µl EB buffer and stored at -80 ℃ until use.

Construction of ChIP-seq libraries. The sequencing libraries were constructed using TruePLEX DNA-seq kit (Rubicon Genomics) or Accel-NGS 2S plus DNA library kit (Swift bioscience). GM12878 samples starting with ≥ 100 cells using DNA oligo linker were

prepared using TruePLEX kit. The other samples are prepared using Accel-NGS 2S plus kit.

1) TruePLEX DNA-seq kit

DNA sample (10 µl) was mixed with 2 µl template preparation buffer and 1 µl template preparation enzyme. The mixture was incubated at 22 ℃ for 25 min, followed by 55 ℃ for 20 min. 1 µl library synthesis buffer and 1 µl library synthesis enzyme were mixed with sample (13 µl) and incubated at 22 ℃ for 40 min. 25 µl library amplification buffer, 1 µl library amplification enzyme, 2.25 µl EVAGreen dye (31000-T, Biotium), 1.75 µl H_2O and 5 µl indexing reagent were mixed with sample (50 µl in total). The sample was amplified at 72 ℃ for 3 min, 85 ℃ for 2 min, 98 ℃ for 2 min, 4 cycles of (98 ℃ for 20 s, 67 ℃ for 20 s and 72 ℃ for 20 s) and certain number of cycles of (98 ℃ for 20 s and 72 ℃ for 50 s) depending on the amount of starting materials (9 cycles for 10,000 cells, 12 cycles for 1000 cells, 13-14 cycles for 500 cells and 15-17 cycles for 100 cells). The sample was purified with 50 µl SPRIselect beads (Beckman Coulter) and eluted to 7 µl EB buffer (Qiagen).

2) Accel-NGS 2S plus DNA library kit

There are five steps in the library preparation procedure, Repair I, Repair II, Ligation I, Ligation II and PCR. After each step, SPRIselect was used to purify sample. Repair I: 10 µl sample was mixed with 43 µl low EDTA TE buffer, 6 µl buffer W1 and 1 µl buffer W2. The mixture was incubated at 37 ℃ for 10 min. The sample (60 µl) was purified by 84 µl SPRIselect beads. Repair II: 30 µl low EDTA TE buffer, 5 µl buffer G1, 13 µl reagent G2, 1 µl enzyme G3 and 1 µl enzyme G4 were mixed with dried beads. The mixture was incubated at 20 ℃ for 20 min and purified by 60 µl SPRIselect beads. Ligation I: 20 µl low

EDTA TE buffer, 3 µl buffer Y1, 2 µl enzyme Y3 and 5 µl reagent Y2 were mixed with dried beads. The mixture was incubated at 25 ℃ for 15 min and purified by 25.5 µl SPRIselect beads. Ligation II: 30 µl low EDTA TE buffer, 5 µl buffer B1, 2 µl reagent B2, 9 µl reagent B3, 1 µl enzyme B4, 2 µl enzyme B5 and 1 µl enzyme B6 were mixed with dried beads. The mixture was incubated at 40 ℃ for 10 min and purified by 42.5 µl SPRIselect beads. The DNA was eluted from beads to 22 µl low EDTA TE buffer. PCR: 2.5 µl EVAGreen dye, 7.5 µl low EDTA TE buffer, 5 µl reagent R1, 4 µl reagent R2, 10 µl reagent R3, 1 µl enzyme R4 are mixed with 20 µl sample. The sample was denatured at 98 ℃ for 30 s and amplified by certain number of cycles of (98 ℃ for 10 s, 60 ℃ for 30 s and 68 ℃ for 60 s) depending on the amount of starting materials (8 cycles for 10,000 cells, 10 cycles for 1000 cells, 13 cycles for 100 cells and 14-16 cycles for <100 cells). After PCR amplification, a double size selection using SPRIselect beads was performed following the product manual. Briefly, large DNA fragments were removed by adding 25 µl beads and incubating for 5 min. The beads that had large DNA bound were discarded and the supernatant was preserved. 12.5 µl AMPure beads were then added into the supernatant and incubated for 5 min. The supernatant was discarded and DNA bound to beads was eluted into 7 µl EB buffer.

Library quality control and quantification. Library fragment size was determined using high sensitivity DNA analysis kit (Agilent) on the TapeStation. The libraries were quantified by KAPA library quantification kit (Kapa Biosystems) and pooled at 10 nM for sequencing by Illumina HiSeq 4000 with single-end 50 nt read. Typically, 20 million reads for H3K4me3 and H3K27ac, and 40 million reads for H3K27me3 were generated per library.

ChIP-qPCR. 25 μl of qPCR (12.5 μl of 2× iQ SYBRGreen Supermix (1708882, Bio-Rad), 1 μl of 10 μM each primer (**Table 5.1**); 95 ℃ for 10 min followed by 45 cycles of 95 ℃ for 15 s, 56 ℃ for 30 s and 72 ℃ for 30 s) was used for GM12878 cells. 20 μl of q PCR (10 μl of 2× iQ SYBRGreen Supermix, 1.4 μl primer mixture (**Table 5.1**, Active Motif); 95 ℃ for 2 min followed by 45 cycles of 95 ℃ for 15 s, 58 ℃ for 20 s and 72 ℃ for 20 s) was used for mouse samples.

The ChIP-qPCR results were presented as relative fold enrichment, which is the ratio of percent input between a positive locus and a negative locus. Percent input was computed using the following equation:

$$Percent\ input = 2^{(Ct_{input} - Ct_{IP})} \times CN_{input} \Big/ CN_{IP} \times 100\%$$

where Ct_{input} and Ct_{IP} are the Ct values of input and ChIPed DNA, respectively; CN_{input} and CN_{input} are the cell number of input and ChIPed sample, respectively.

ChIP-Seq reads mapping and normalization. Sequencing reads were trimmed by trim galore! with default settings. The trimmed reads were aligned to hg19 or mm9 genome by and bowtie [181]. Mapped reads were used for narrow peak calling with MACS2 (q<0.05) [199] and broad peak calling with SICERpy [200] (-rt 0 --windowSize 1000 --gapSize 3). Uniquely mapped reads from both ChIPed and input samples were extended to 250 bp and used to compute a normalized signal for each 100 nt bin across the genome. Normalized signal was defined as following:

$$normalized\ signals = IP\left(\frac{Reads\ in\ each\ bin}{Total\ uniquely\ mapped\ reads} \times 1000000\right) -$$

$$Input\left(\frac{Reads\ in\ each\ bin}{Total\ uniquely\ mapped\ reads} \times 1000000\right).$$

Construction of Receiver Operating Characteristic (ROC) curves. We compared the performance of SurfaceChIP-Seq to that of the state-of-the-art method MOWChIP-seq [29] using ROC curves. We focused on promoter regions (defined as 2000 bp upstream and 500 bp downstream of a transcription start site (TSS) based on RefSeq genes). The ROC curves were constructed using previously described methods [29]. The gold-standard data were generated by the ENCODE consortium (GEO: GSM733708 and GSM945188). The ROC curves of MOWChIP-seq are slightly different from the original paper due to the selection of peak calling approach.

Correlation analysis and of SurfaceChIP-Seq data. Normalized ChIP-Seq signals in all promoter regions in the genome were extracted. Promoter regions were defined as ± 2kb around TSS. Averaged signals across the promoter region was used for computing Pearson correlation coefficient.

Differentially binding promoters. Unnormalized ChIP-seq signals in all promoter regions (± 2kb around TSS) in the genome were extracted and analyzed by DEseq2 [201].

Construction of RNA-seq libraries. Freshly extracted nuclei mixture or sorted nuclei were immediately used for total RNA extraction using RNeasy Mini Kit with additional DNase treatment step to remove gDNA contamination. RNA-seq libraries were prepared using

SMART-seq v4 Ultra Low Input RNA kit and Nextera XT DNA library prep kit following instructions. Briefly, ~1 ng total RNA (9.5 µl) was mixed with 1 µl of 10x reaction buffer. The sample was mixed with 2 µl of 3' SMART-seq CDS primer II A and incubated at 72 ℃ for 3 min. The sample was placed on ice for another 2 min. 5.5 µl of Master Mix and 2 µl of SMARTScribe Reverse Transcriptase were added into sample (20 µl in total). The sample was incubated at 42 ℃ for 90 min and 70 ℃ for 10 min to synthesize first-strand cDNA. The cDNA was amplified in 14 cycles of PCR. The PCR product was purified by AMPure beads. 100-150 pg purified cDNA was used for Nextera library preparation. The cDNA was tagmented at 55 ℃ for 5 min and neutralized at room temperature for another 5 min. The tagmented DNA was amplified in 12 cycles of PCR and purified by 0.6x AMPure XP beads. The fragment size of the library was determined by a Tapestation and quantified by KAPA qPCR library quantification kit. Each library was pooled at 10 nM for sequencing by Illumina HiSeq 4000 with single-end 50 nt read.

RNA-seq data analysis. Sequencing reads were trimmed by trim galore! with default settings. The trimmed reads were aligned by Tophat [186]. The aligned reads were further analyzed by cufflinks [187]. The gene expression level was plotted by cummerbund (http://compbio.mit.edu/cummeRbund/), Seqmonk and custom R scripts. Differentially expressed genes were identified by cummerbund.

Bivalent promoters. The bivalent promoters were defined as the promoter regions that overlapped with both H3K4me3 and H3K27me3 peaks. Promoter regions were defined as ± 2kb around TSS.

Prediction of enhancers and super enhancers. Enhancers were predicted using the CSI-ANN [202] and normalized SurfaceChIP-Seq signals (H3K4me3, H3K27ac and H3K37me3) across the genome. The predicted enhancers and H3K27ac SurfaceChIP-Seq data were then used as the input to predict super enhancers using the ROSE (http://bitbucket.org/young_computation/rose). Enhancers within 15000 bp were stitched together. Enhancers located within ± 2000 bp from annotated TSSs were excluded. Genes closest to the super enhancers were extracted as targets for gene ontology (GO) term enrichment analysis using DAVID [185].

Transcription factor motif enrichment analysis. We used the coordinates of the predicted enhancers as input for motif analysis. We used Homer [203] (-size 200 -mask -nomotif) to find the enrichment (p value) of the known transcription factors in the default database.

COMSOL modelling of chromatin capture efficiency. The time-dependent diffusion process in the microfluidic device was modelled using COMSOL Multiphysics 4.3 [12]. We used 2×10^{-11} for diffusivity of 170-bp chromatin and 5×10^5 for association rate constant of chromatin-antibody binding [204].

5.4 Conclusion

In summary, we demonstrated that SurfaceChIP-Seq with as few as 30-100 cells generated high quality epigenomic profiles. Our microfluidic platform allows running up to 8 different reactions in parallel which dramatically improved the throughput. SurfaceChIP-

seq may potentially complement other high-sensitivity ChIP technologies [29, 31, 38, 192, 205]

and achieve higher sensitivity and coverage.

CHAPTER 6 – RNA extraction from a mycobacterium under ultrahigh electric field intensity in a microfluidic device

6.1 Introduction

Profiling transcriptomes (the set of all RNA molecules) is critical for understanding the functional elements of the genome and disease processes [206]. Various technologies have been developed in recent years, such as real time PCR [207], microarrays [208] and RNA sequencing (RNA-seq) [206, 209, 210], to detect and quantify mRNA for understanding physiological events. The purity and integrity of input RNA are critical for the success of these RNA based analysis. Compromise in RNA quality leads to variable results [211, 212]. There is a growing demand for mRNA extraction methods that allow transcriptomic profiling of all species.

Mycobacteria are nonmotile, aerobic and acid-fast bacteria, including highly pathogenic species that cause tuberculosis and leprosy [213]. Compared to other bacteria, mycobacteria have thick cell wall that is hydrophobic, waxy, and rich in mycolic acids/mycolates. Analysis of intracellular contents from mycobacteria is challenging due to this structural characteristic.

Several methods have been developed for RNA isolation from bacteria. Chemical disruption, including TRIzol-based method [214] and hot-phenol based method [215], is traditionally used for bacteria RNA extraction. However, the procedures are usually tedious and time-consuming, taking several hours to a few days [216, 217]. The chemicals (SDS and phenol) involved in these procedures often cause RNA fragmentation and

result in variability in RNA quality and analysis results[217]. Bead beating is considered to be the state of the art for RNA extraction from lysis-resistant bacterial cells including mycobacteria [218]. TRIzol is often added in bead beating to improve RNA stability and facilitate cell wall disruption (via denaturing proteins and inhibiting RNases) [219]. Bacterial cells and beads are contained in a closed tube and then subject to high-frequency oscillation. The high shear stress generated by periodic vertical flow leads to mechanical lysis. This procedure is typically applied to a large number of cells ($>10^8$ cells). This creates difficulty for studying slow-growing mycobacteria and probing a low number of cells [220]. Such release tends to be incomplete. The mechanical mechanism is also hard to replicate on a microfluidic device.

Electric lysis is a rapid physical method for cell membrane disruption and intracellular content release [13, 221]. Electric lysis typically refers to irreversible cell electroporation that occurs under application of electrical pulses with defined intensity and duration. When the field intensity and duration of these electrical pulses exceed certain threshold (that is specific to the cell type), cells are irreversibly lysed and intracellular molecules are released into surrounding solution. Electric lysis can be completed rapidly (within seconds to minutes) and does not involve the use of chemical/biological reagents that may potentially interfere with downstream assays. Although electroporation has been extensively utilized for releasing intracellular proteins [13, 221, 222], there has been very little work on using it to release nucleic acids in general [83, 125, 223]. None of the previous works examined the effectiveness of electric lysis on mycobacteria that are generally considered highly resistant to most lysis methods.

Here we describe a rapid mRNA extraction from *Mycobacterium smegmatis* under ultrahigh-intensity (up to 8000 V/cm) electric lysis on a microfluidic device [86]. We formed a packed bed of microscale silica beads in the device to trap the mycobacterial cells. Electric pulses were then applied to electrically lyse trapped *M. smegmatis* within 3 minutes. Cell lysate was directly used for qRT-PCR analysis without further treatment. We show that our mRNA extraction efficiency was 10-20 times higher than bead beating.

6.2 Materials and Methods

Microfluidic chip fabrication and operation

Two-layered PDMS chip was fabricated by multilayer soft lithography [5, 147]. Photomasks were designed by Freehand MX (Macromedia, San Francisco, CA) and printed on transparency film at 4000 dpi resolution. The fluidic layer master was fabricated in SU-8 2002 (Microchem, Newton, MA) and AZ 9260 (Clariant, Charlotte, NC) with the thickness being 2 and 10 µm, respectively. Micro-pillars were placed in fluidic chambers to avoid collapse. The master was heated at 130 ℃ for 30 s to form rounded cross-sectional profile for the features in AZ 9260. The control layer master was fabricated in SU-2025 with 24 µm thickness. The control layer was made by spinning PDMS prepolymers (RTV615A: RTV615B=20:1, R. S. Hughes, Sunnyvale, CA) at 500 rpm for 10 s and then at 1700 rpm for 30s. This resulted in a thickness of 67 µm for the control layer PDMS. The fluidic layer had a composition of RTV615A: RTV615B =5:1 and a thickness of ~0.4 cm. Both layers of PDMS were cured at 80℃ for 15 min. The two layers were then aligned, brought into contact and baked for 1 h at 80℃. The two-layered PDMS structure was then peeled off from the control layer master and access holes were

punched. Glass slides (VWR, Radnor, PA) were cleaned in a basic solution (H_2O: 27%

NH_4OH: 30% H_2O_2 = 5:1:1, volumetric ratio) at 80 ℃ for 1 h, rinsed with ultrapure water

and blown dry using nitrogen. Finally, the PDMS structure was bonded to the pre-cleaned

glass slide after plasma oxidation of both surfaces (Harrick Plasma, Ithaca, NY) and

baked at 80℃ for 1 h for further strengthening of the device.

We wrote two LabVIEW (National Instruments, Austin, TX) programs to operate

on-chip pneumatic valves and generate electric pulses with designated duration/direction,

all via a DAQ card (NI SCB-68 from National Instruments) [20,24,25]. One program controlled

actuation of on-chip pneumatic valves (under a pressure of 25 psi) via operation of

solenoid valves (18801003-12VDC, ASCO Scientific, Florham Park, NJ). The other

defined the pulse duration and direction. Two platinum electrodes were inserted into the

reservoirs at inlet 2 and the outlet with direct contact to the solution (**Fig. 6.1**). The two

electrodes were connected to DC voltage provided by a high voltage power supply

(PS350/5000V-25W, Stanford Research System, Sunnyvale, CA) via an H-bridge circuit.

The H-bridge circuit consisted of 4 high voltage reed relays (5501-05-1, Coto Technology,

Kingstown, RI). Two of the relays worked together (by switching closed and then open at

the same time) to generate a pulse of defined duration in one direction (while the other

two relays kept open). On the other hand, when the other two relays were actuated in the

same fashion, a pulse in the opposite direction was generated. To eliminate potential pH

changes in the electroporation solution due to electrolysis of water under long pulses, we

applied a series of pulses by changing the current direction after each pulse. In this study,

each individual pulse was 5 s in duration and there was a 2 s interval between pulses.
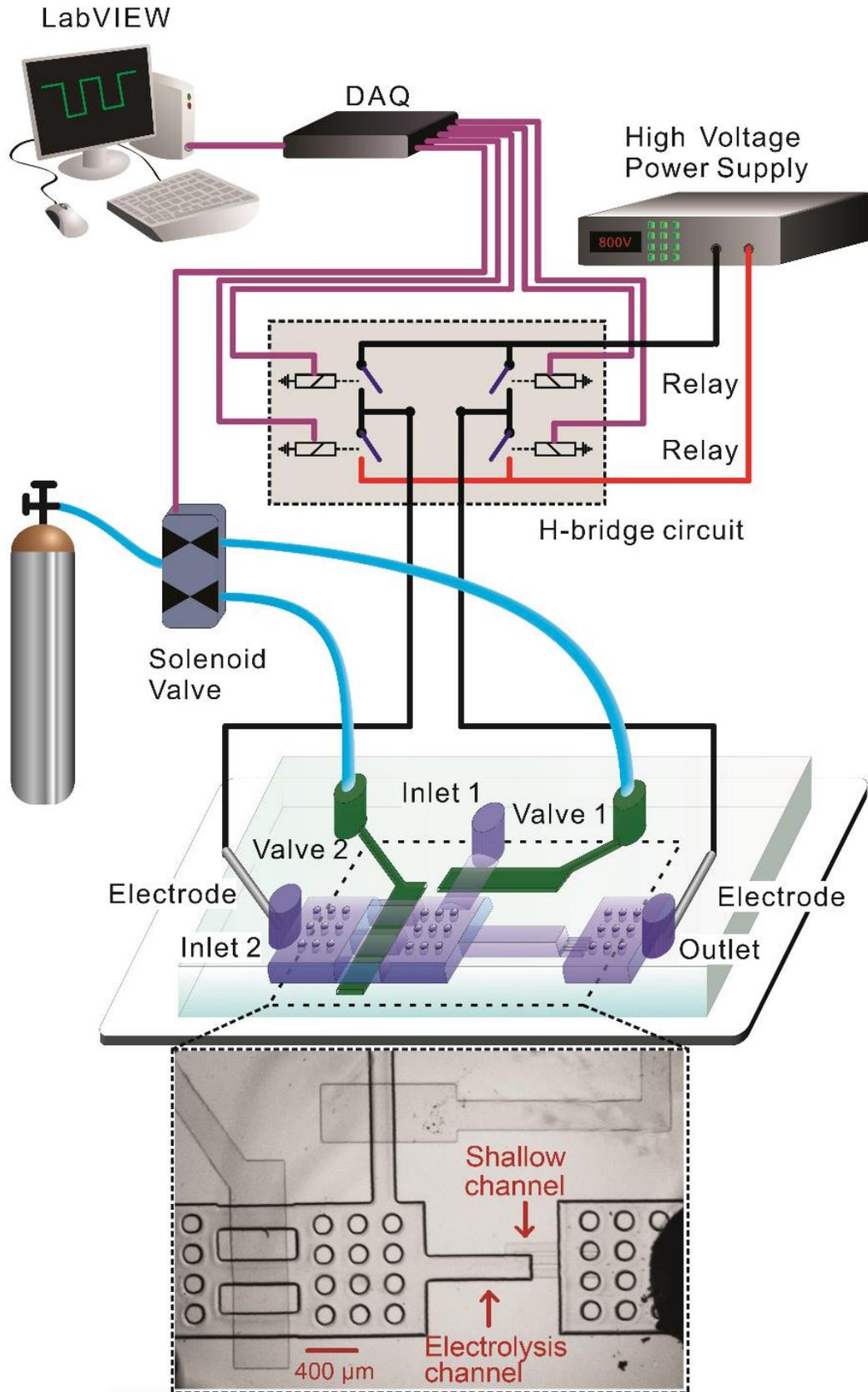
Fig. 6.1 Experimental setup for conducting ultrahigh-intensity electrolysis of M. smegmatis. Electric pulses were applied across a PDMS microfluidic channel via two platinum electrodes. Pulses were generated by cutting off constant voltage generated by a power supply. The pulse duration and direction were set in a LabVIEW program that operated a H-bridge circuit (consisting of 4 high-voltage relays) via a data acquisition (DAQ) card. The two-layered microfluidic device contained two pneumatic valves. The dimensions of the electrolysis channel were 800 µm × 200 µm × 10 µm. There were circular supporting pillars in the large chambers to prevent collapsing.

Cell culture

Mycobacterium smegmatis MC2 155 expressing Green Fluorescent Protein (GFP) was cultured in Lysogeny Broth (LB) with 0.05% Tween 80 and 100 µg/µl hygromycin at 37℃ under continuous shaking at 200 rpm. The cell culture was centrifuged at 3000 rpm for 10 min and resuspended in H_2O. The concentration of the cell culture was estimated by McFarland standards.

RNA extraction by bead beating

The cell culture was centrifuged at 4000 rpm for 10 min. 5×10^9 to 10^{10} cells (with the actual value calculated based on optical density reading taken with a spectrophotometer) were then resuspended in 1 ml TRIzol (Life Technologies). The mixture was then transferred into a lysing matrix B bead beating tube (MP Biochemicals, Santa Ana, CA) on ice. 300 µl chloroform was added and the bead beating tube was shaken vigorously. Bead beating using maximum power was conducted twice (45 seconds the first time and 30 seconds the second time), with 5-minute rest time in

between. The mixture was centrifuged at 10000 rpm for 15 min at 4 ℃ and the top aqueous layer was collected. The sample was mixed with 600 µl isopropanol and stored at -20 ℃ overnight. The RNA sample was centrifuged at 10000 rpm for 10 min at 4 ℃, washed with 75% ethanol and resuspended in H_2O.

Quantitative reverse transcription PCR (qRT-PCR) sssay

Cell lysate was eluted from the microfluidic device by 20 µl RNase-free water and then mixed with 1x RNAsecure Reagent (Life Technologies) to deactivate RNases and protect the integrity of RNA. The mixture was incubated at 60 ℃ for 10 min, then cooled to 4 ℃. The samples were directly used in qRT-PCR (four pairs of primers are listed in **Table 6.1**) or stored at -20 ℃ until use. QPCR assays were performed in 20 µl aliquots in a Real-Time PCR Detection System (CFX Connect, BIO-RAD, Hercules, CA). Power SYBR Green RNA-to-CT™ 1-Step Kit (Life Technologies) was used to measure mRNA copy numbers following the manufacturer's protocol. A pair of primers (100 nM each) were added into each aliquot. The mixture was incubated at 48 ℃ for 30 min for reverse transcription. The mixture was then denatured at 95 ℃ for 10 min, followed by 40 cycles of 95 ℃ for 15 s and 60 ℃ for 60 s for DNA amplification. No-template controls (NTCs) were included to monitor non-specific amplification.

Table 6.1 The primer sequences used in qRT-PCR.

mRNA	Forward primer sequences	Reverse primer sequences
SigA	GACTACACCAAGGGCTACAAG	TTGATCACCTCGACCATGTG
ClpX	TGACCAAGTCCAACATCCTG	CAACATCCTCACCGACGTAG
DnaE	CGTCTACCAAGAGCAGATCATG	TTCCTTGAAGCCCTTGTACTC
KatG	GACCGCGAATGACCTTGTGT	TGTCGGACTGGGCATAAACC

RNA standard synthesis and quantification

We established a standard curve to link the copy number of mRNA with qPCR signal (Ct value). *In vitro* transcribed RNA was used as the template to generate the standard curve. Each of the four pairs of primers in **Table 6.1** with added T7 promoter sequence (5' - TAA TAC GAC TCA CTA TAG GG - 3') was used to amplify a specific mRNA (from total RNA extracted by bead beating) by RT-PCR for 40 cycles. The size of RT-PCR products were verified by gel electrophoresis (**Fig. 6.2**). The amplification product (DNA) contained the T7 promotor sequence on both ends that could be recognized by T7 phage RNA polymerases. About 75 ng amplification product was mixed with nucleotides (0.5 mM in final concentration), T7 polymerase and 1x Transcription Buffer polymerases (MAXIscript T7 *in vitro* transcription kit, Life Technologies) and incubated at 37 ℃ for 1 h. TURBO DNase 1 was added and incubated at 37 ℃ for 15 min to digest template DNA. 25 mM EDTA was used to stop digestion and block heat-induced RNA degradation. The produced RNA was purified twice by precipitating with ammonium acetate/ethanol to remove unincorporated NTPs and DNase. The copy number of the RNA was measured by Qubit 2.0 Fluorometer with Qubit RNA HS Assay Kit (Life Technologies). The RNA was serially diluted to generate samples for the standard curve using qRT-PCR.

COMSOL modelling

COMSOL Multiphysics 4.3 (Burlington, MA) was used to simulate the electric field intensity in the microfluidic chamber. The "Electric Current (ec)" module was used to study the steady state of model. The 3-D model was based on the actual geometry of the microfluidic structures with the exception of the packed bed. Because the resistivity of silica beads (1×10^{13} Ωm) is much higher than that of water (2×10^5 Ωm), we treated
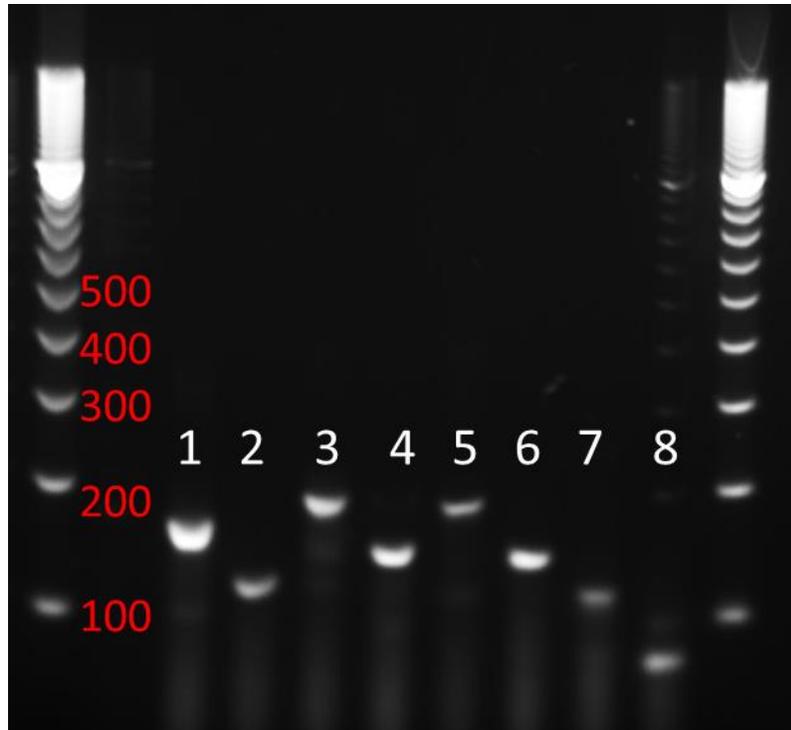
Fig. 6.2 RT-PCR product using the primers in Table 6.1 and the primers containing added T7 promoter sequences. The RT-PCR products generated using primers containing T7 promoter sequences were 40 bp longer than the DNA templates generated using primers in Table 6.1. PCR products and 100 bp DNA ladder (Biorad) were loaded on 2% agarose gel. The PCR products containing T7 promoter sequences were in Lane 1 (SigA), 3 (ClpX), 5 (DnaE) and 7 (KatG). The PCR products without T7 promoter sequence (using primers in Table 6.1) were in Lane 2 (SigA), 4 (ClpX), 6 (DnaE) and 8 (KatG).

silica beads as electrically insulating. Thus the segment that contained the packed bed

was treated as having a cross-sectional area smaller than that of the channel (with the equivalent cross-sectional area occupied by the insulating beads deducted). Water was used in all other domains with electrical conductivity of 5.5×10^{-6} S/m. To simplify the modeling, we did not consider gas bubbles and treated the solution as static.

The modeling is based on the governing equation:

$$\frac{\partial (\nabla \cdot D)}{\partial t} = -\nabla \cdot J \qquad (1)$$

where D is electric flux density and J is the electric current density.

The electric current density is expressed by:

$$J = \sigma E \qquad (2)$$

where σ is the electrical conductivity and E is electric field. The electric flux density (D) is also expressed by:

$$D = \epsilon_0 \epsilon_r E \qquad (3)$$

where ϵ_0 is the vacuum permeability and ϵ_r is relative permeability. The system was electrically insulated on all boundaries. 786 V across the two reservoirs was applied.
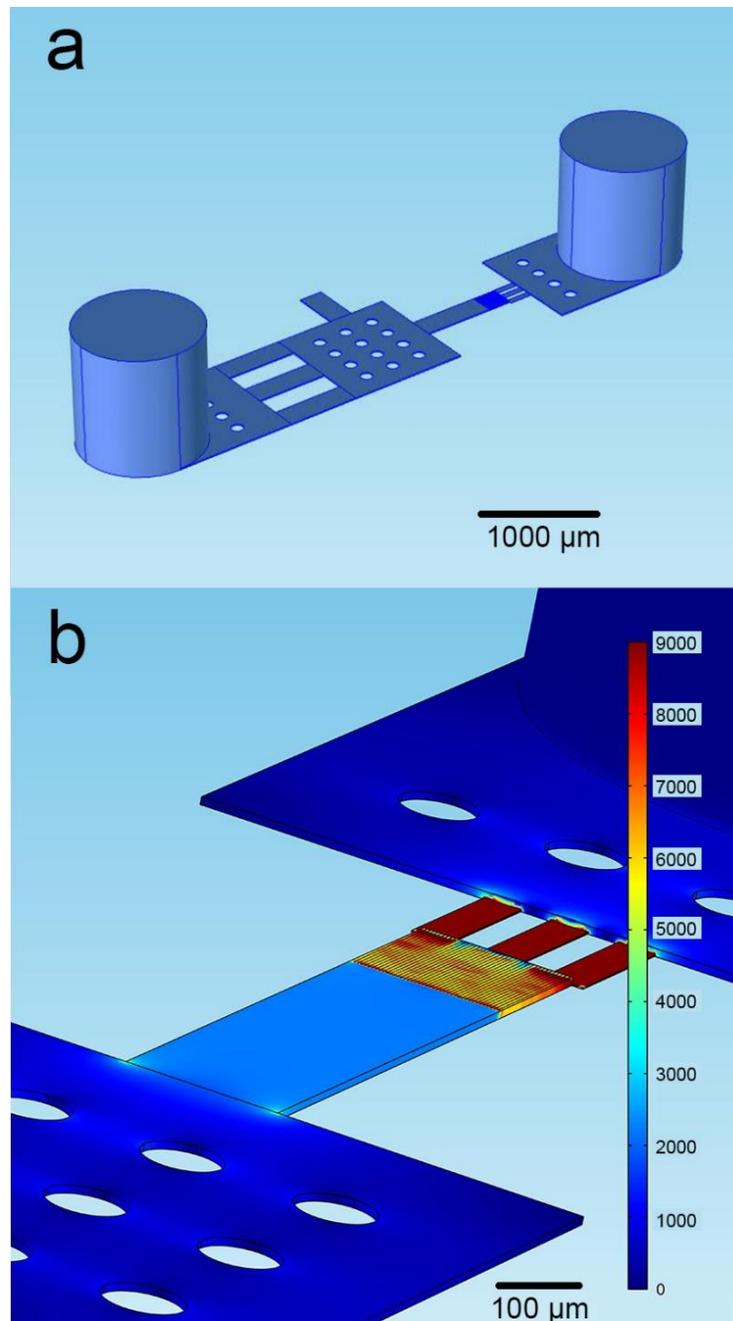
Fig. 6.3 COMSOL Multiphysics modelling of electric field intensity in the electrolysis channel with packed beads. (a) The overview of the modelled structure. (b) A zoom-in image of the electrolysis channel with 786 V DC voltage applied across the two reservoirs. The average field intensity in the space among beads was 6000 V/cm.

6.3 Results and Discussion

Device and protocol for microfluidic electric lysis.

We designed a two-layered microfluidic electric lysis chip (**Fig. 6.1**). There were two inlets and one outlet. The depth of channels and chambers in the fluidic layer was 10 µm with the exception of the shallow channels (having a depth of 2 µm). Prior to the experiment, the device was primed by flowing water in via both inlets, removing debris and bubbles. Microscale silica beads (d~4.8 µm, SS05N, Bangs Laboratories, Fishers, IN) and mycobacterial cells (*M. smegmatis*) then flowed into the device via inlet 1, while keeping valve 2 closed. 20 µl mixture of beads and cells (~2 × 10^7 cell/ml) was packed against the downstream shallow channel to form a packed bed in the electrolysis channel with mycobacterial cells trapped inside. In our experiment, roughly 2×10^5 cells were captured in the electrolysis channel each time and loading of cells and beads was stopped by closing valve 1. Two platinum (Pt) electrodes were inserted in the inlet 2 and the outlet for application of electric pulses. Valve 1 was closed and there was a flow of 1 µl/min coming in from inlet 2 during the application of the electric pulses.

Compared to electric lysis of bacterial cells such as *E. coli* [221, 224], electric lysis of *M. smegmatis* required significantly longer pulse duration and higher field intensity. Several features of our device and operation were designed to accommodate these requirements. First, we found that even low-conductivity buffer (e.g. 1 mM $MgSO_4$, 8 mM Na_2HPO_4, 2 mM KH_2PO_4 and 250 mM sucrose) used in our previous works [12, 15, 225, 226] generated too many bubbles when the field intensity was >2000 V/cm). Thus we flowed ultrapure water (RNase-free) during the application electric pulses. Second, we applied a flow of 1 µl/min (coming from inlet 2) during application of electric pulses. The flow carried

away generated heat and bubbles during the operation. Finally, we programmed the LabVIEW code and designed an H-bridge circuit (including four relays) so that the direction of the electric field was alternated after each pulse (of 5s duration). This was to prevent significant local pH change due to electrolysis of water associated with long pulses.
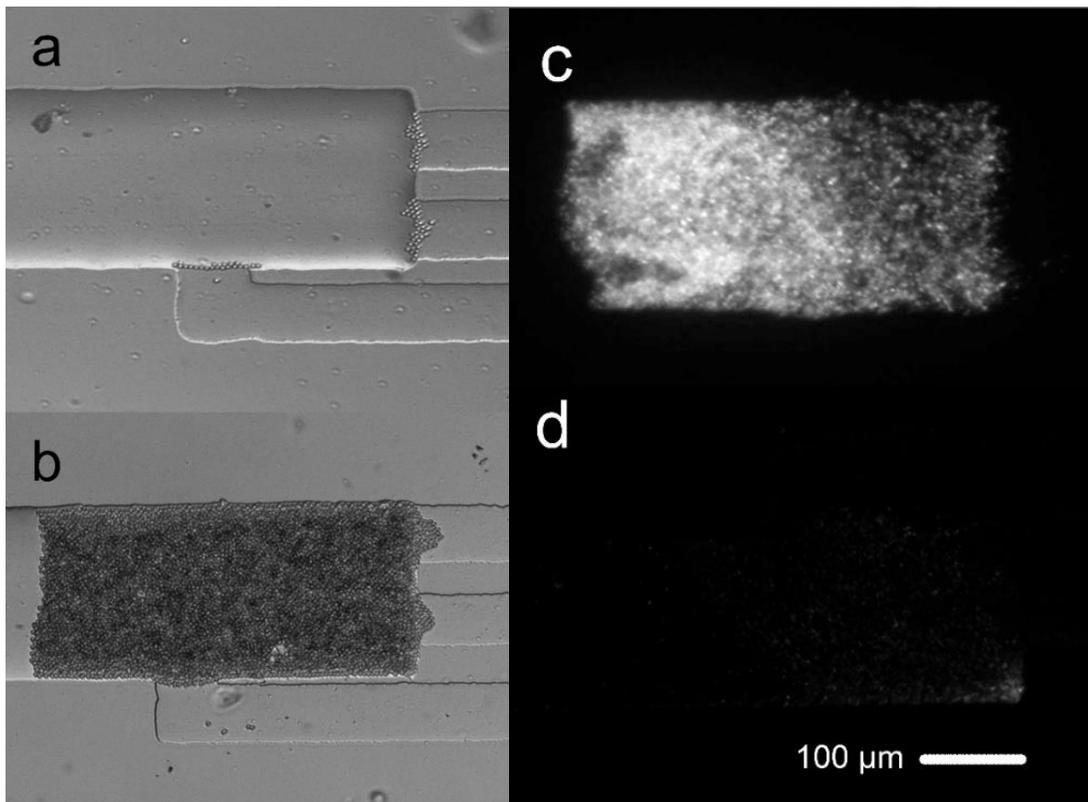
Fig. 6.4 The procedure for microfluidic electrolysis of *M. smegmatis*. (a) Bright field image shows that beads were stopped by the shallow channel and started to accumulate in the electrolysis channel. (b) Bright field image shows the formation of packed bed of beads/ M. smegmatis cells. (c) Fluorescence image shows the GFP expressing M. smegmatis cells before electrolysis. The packed bed roughly contained 2×105 cells. (d) Fluorescence image after electrolysis (30 Pulses of 6000 V/cm and 5 s each).

Other than the channel dimensions and the voltage applied, the field intensity of the pulses was affected by the presence of nonconductive beads. There was a substantially higher resistance associated with the packed bed of beads than that of the same channel segment without the beads. We used COMSOL Multiphysics to model the field intensity within the packed bed (**Fig. 6.3**). We show that the presence of the packed bed significantly increased the local field intensity (by ~3x) (compared to the channel segment without beads).

We used a *M. smegmatis* strain that expressed green fluorescent protein (GFP). In **Fig. 6.4**, we show the process of bead/cell packing and the electrolysis-based release of intracellular materials. The initial stopping of beads at the interface of the shallow channel and the electrolysis channel (**Fig. 6.4a**) led to the formation of a packed bed with mycobacterial cells trapped in the gaps among beads (**Fig. 6.4b**). The fluorescence emitted from the cells was observed in **Fig. 6.4c** before electrolysis and disappeared after electrolysis by 20 5 s pulses with a field intensity of 6000 V/cm (**Fig. 6.4d**), suggesting release of intracellular molecules.
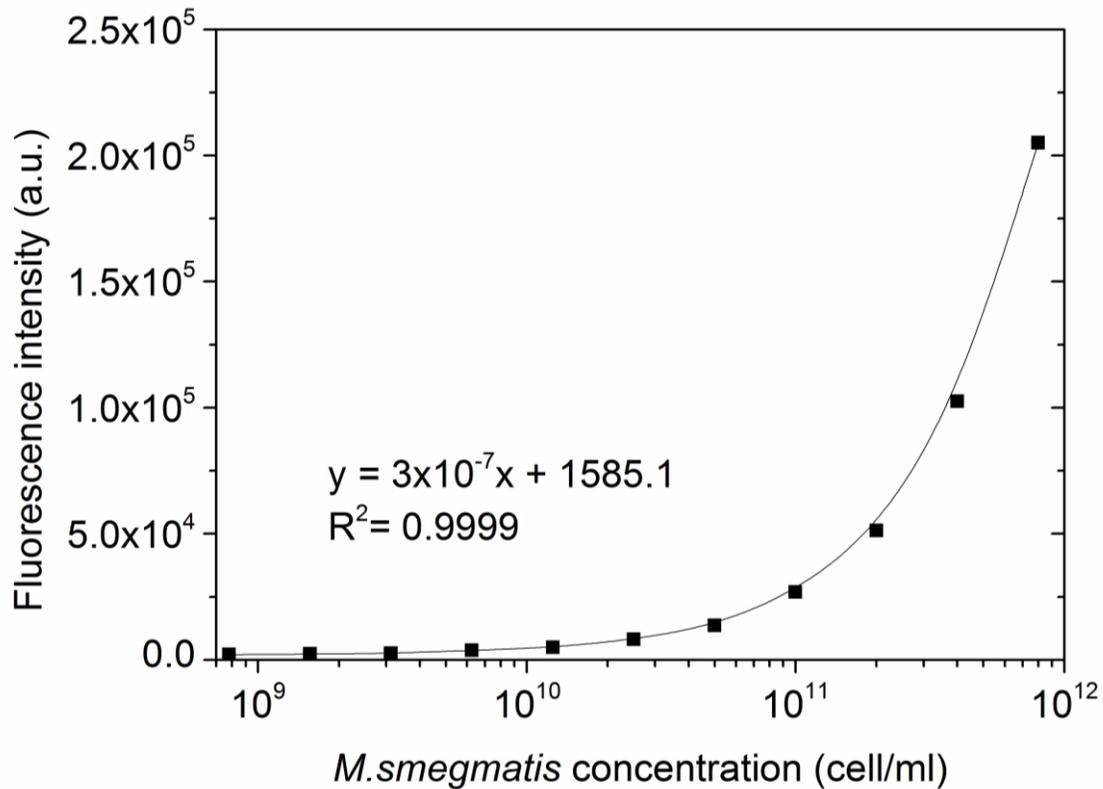
Fig. 6.5 A standard curve for quantifying mycobacterium concentrations in the electrolysis channel. GFP expressing M. smegmatis with different concentrations were loaded in a straight microfluidic channel with the same depth as the electrolysis channel. The fluorescence intensity was calculated based on z-stacking of 10 images.

We quantified the amount of RNA release by electrolysis using quantitative reverse transcription polymerase chain reaction (qRT-PCR). The RNA extraction was quantified by

$$mRNA\ copy\ per\ cell = \frac{10^{\left(\frac{intercept-Ct}{slope}\right)} \times 20\ \mu l}{\frac{(FI - 1581.1)}{3.0 \times 10^{-7}} \times V} \times (1 - DNA\%) \quad (4)$$

where 1581.1 and 3.0×10^{-7} were parameters extracted from the working curve in **Fig. 6.5**. Intercept and slope were calculated for each mRNA in **Fig. 6.6**. Ct was the RT-qPCR signal. The number of GFP-expressing mycobacterial cells trapped in the packed bed was estimated based on the amount of emitted fluorescence. The fluorescence intensity (FI) was the average of 10 z-stacking images that covered the entire depth of the packed bed (taken after the GFP-expressing mycobacterial cells occupied the packed bed). DNA% was the fraction of a mRNA copy per cell value contributed by DNA (**Fig. 6.7**). Packed bed volume (V) was calculated by

$$V = length\ of\ packed\ bed\ \times 200\ \mu m \times 10\ \mu m \qquad (5)$$
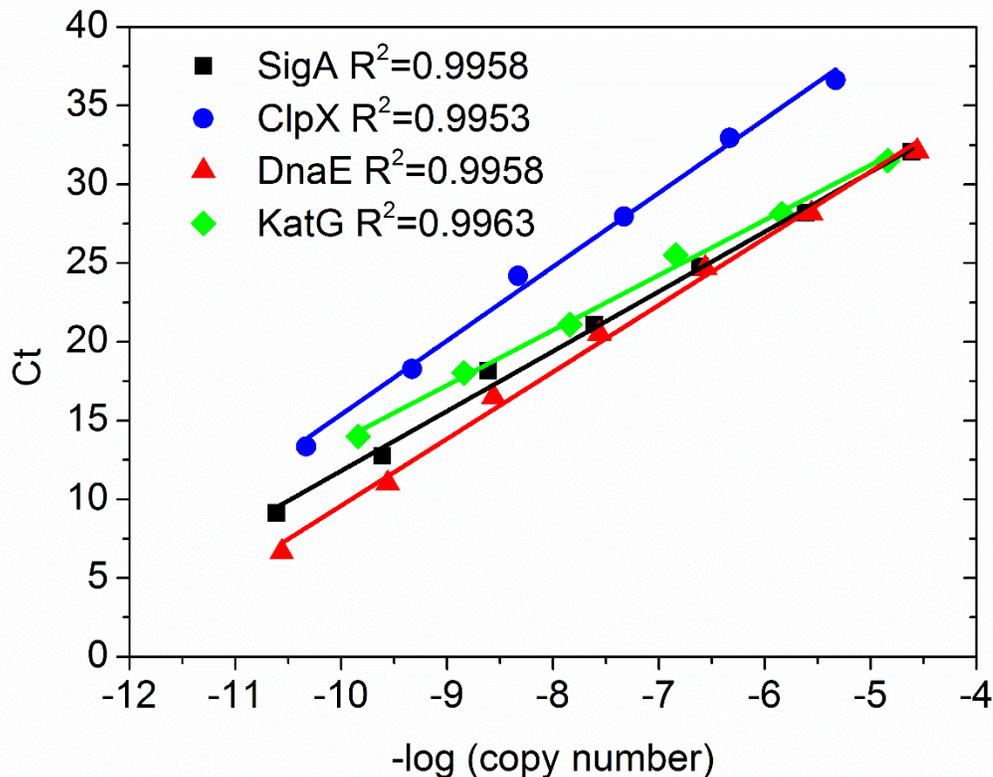
Fig. 6.6 The qPCR standard curves for all 4 mRNAs established using in-vitro synthesized RNA with known copy numbers (obtained by a Qubit 2.0 fluorometer).
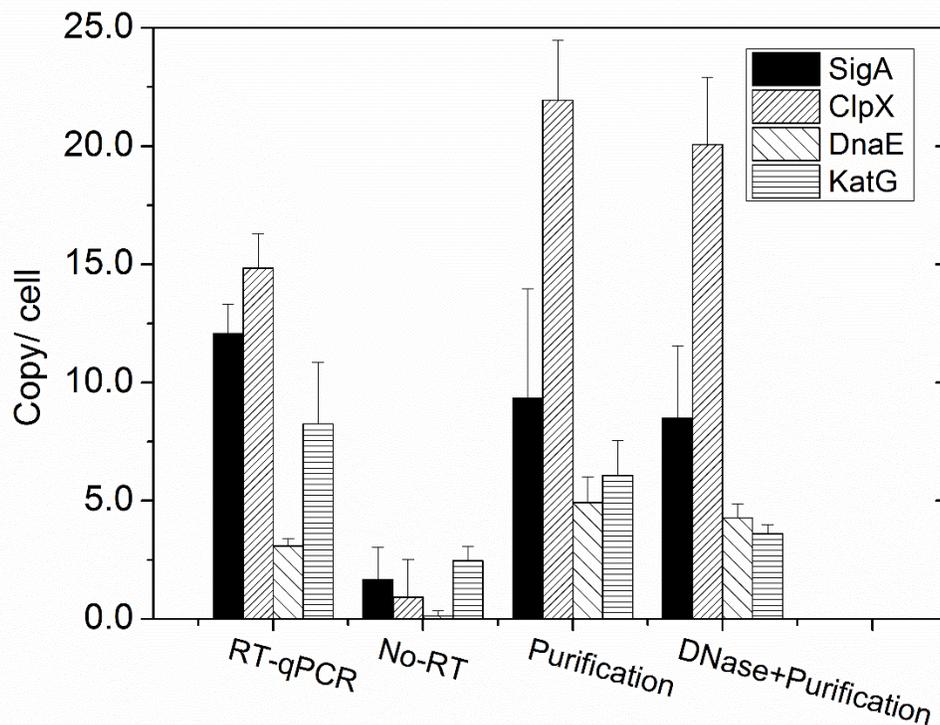
Fig. 6.7 The effect of DNA templates in the lysate on RT-qPCR results. All samples

were extracted by applying 30 pulses of 6000 V/cm at 5 s duration from roughly 2×10^5

M. smegmatis cells (N=3). "No-RT" was performed without reverse transcription step.

"Purification " was performed by purifying the sample using phenol/chloroform

extraction, ethanol precipitation and resuspended in 20 µl water. "DNase+Purification"

was performed by first treating the sample with 0.02 U/µl DNase I at 37 C° for 10 min,

stopped by adding EDTA (at a concentration of 5 mM) and incubated at 75 C° for 10

min. The sample was then purified by phenol/chloroform extraction and ethanol

precipitation. Based on comparisons between RT-qPCR and No-RT, we conclude that

the fractions of RNA copy per cell values generated by DNA templates in the lysate are

14% for SigA, 6% for ClpX, 4% for DnaE and 30% for KatG.

RNA extraction efficiency under various electric lysis conditions.

We characterized RNA extraction efficiency by the electrolysis protocol under various pulse numbers (2 to 30 pulses, all at 6000 V/cm) and field intensities (4000-8000 V/cm, all with 30 pulses) (**Fig. 6.8**). We quantified the copy number per cell for 4 representative mRNAs in the cell lysate under various electric lysis conditions and compared them with those with bead beating (i.e. the conventional lysis method). The number of cells in the electrolysis channel was quantified by fluorescence, as shown in **Fig. 6.5.** RNA quantification was based on standard curves established using *in vitro* transcribed RNAs, which offered more accuracy than using bulk RNA, cDNA or DNA based quantification [227]. **Fig. 6.8a** shows that there was no substantial release of mRNA (under 2 or 4 pulses at 6000 V/cm) until the number of pulses increased to 10. The amount of mRNA release increased with higher number of pulses (from 10 to 20 and 30). Some genes experienced drastically increase with the number of pulses (e.g. SigA from 6.3 copies/cell at 10 pulses to 16.9 and 20.7 copies/cell at 20 and 30 pulses, respectively and ClpX from 3.0 copies/cell at 10 pulses to 12.6 and 18.7 copies/cell at 20 and 30 pulses). The results suggest that a significant number of pulses (or total electrolysis duration) was required to create sufficient cell wall damage for *M. smegmatis* and complete release of intracellular RNAs. Overall, electrolysis provided significantly more complete release of intracellular mRNAs than the current state of the art (i.e. bead beating). We release a factor of 18 and 10 more RNA molecules with 30 pulses than bead beating, for SigA and ClpX, respectively. Furthermore, we also examined the effect of the pulse intensity. As expected, increased field intensity of the pulses from 4000 to 6000 V/cm increased the mRNA release significantly for all genes examined (when the number of pulses was kept

at 30). However, further increase from 6000 to 8000 V/cm did not produce significant additional increase. This indicates that with 30 pulses at 6000 V/cm, the release of mRNA was fairly complete. Further increase in the field intensity or duration did not produce better results and could potentially lead to loss of integrity and function for mRNA due to heating.

Our results demonstrate that electrolysis of hard-to-lyse mycobacterial cells can be conducted with high efficiency under ultrahigh field intensity in a microfluidic device. Our microfluidic lysis protocol takes less than 30 min for each sample (including cell capture, lysis and elution) and may potentially be run in parallel for high throughput of samples. Our microfluidic platform offers advantages for handling small quantities of cells and integration with downstream assays.
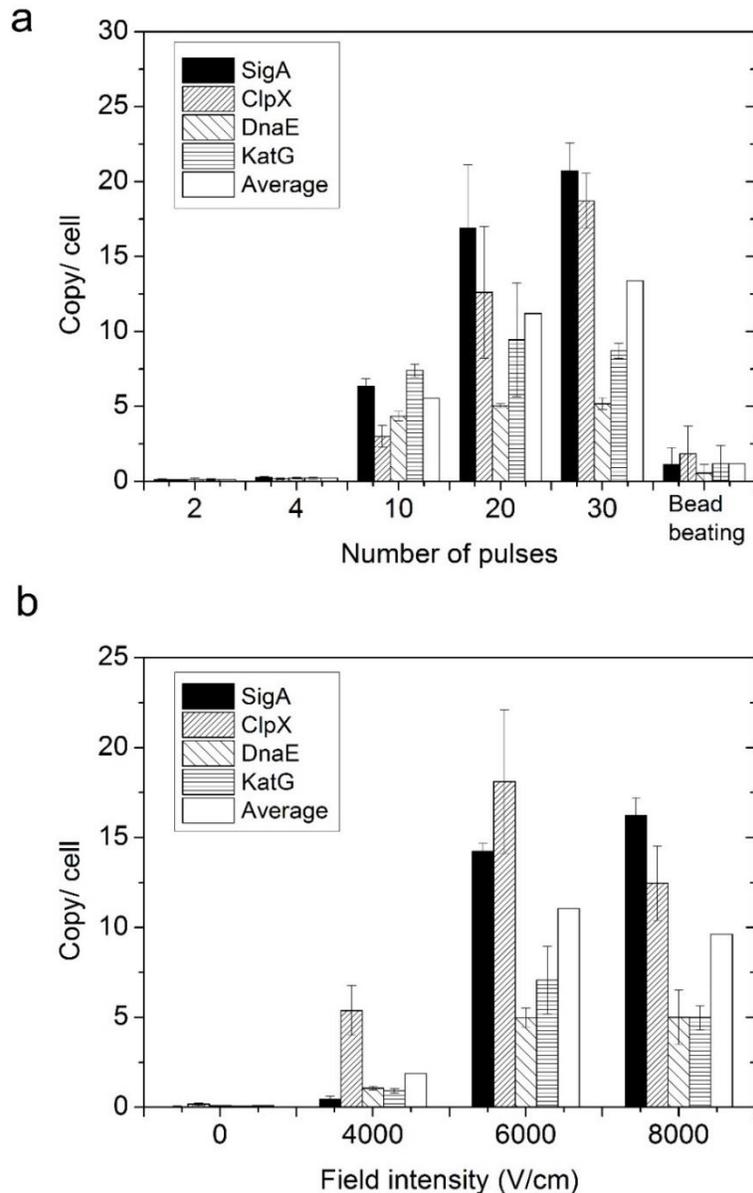
Fig. 6.8 The release of mRNA under various conditions for *M. smegmatis*. The released copy number per cell for 4 mRNAs were quantified using qRT-PCR (n=3). The copy per cell values for microfluidics-produced samples were calibrated by deducting the fraction generated by DNA templates in the cell lysate (Fig. 6.5). (a) The effect of pulse numbers and comparison to bead beating. Various numbers of pulses (5s duration for each pulse with 6000 V/cm intensity) were applied. (b) The effect of pulse intensity. In each case, 30 pulses (5 s each) were applied.

CHAPTER 7 – Electroporation-based delivery of cell-penetrating peptide conjugates of peptide nucleic acids for antisense inhibition of intracellular bacteria

7.1 Introduction

Cell penetrating peptides (CPPs) are short (<30-35 amino acid residues), water-soluble peptides with a net positive charges at physiological pH [228, 229]. CPPs may serve as delivery vectors for large molecules and penetrate cell membranes at micromolar concentrations without causing serious damage to cells. There have been a number of CPPs discovered over the years, including TAT peptide derived from HIV-1 [230-232], transportan and its analogue TP-10 [233-237], penetratin [238] and pVEC [235, 239, 240]. The mechanisms for CPP entry into cells are still in debate. In general, it is believed that most CPPs enter cells via endocytosis while direct penetration may also be involved or even important in certain cases [229]. These mechanisms also may vary based on the experimental conditions and the properties of the cargos [229].

There has been growing interest in using CPPs for delivery of antisense reagents such as siRNAs (small interfering RNAs) and PNAs (peptide nucleic acids, gene-specific artificial oligonucleotide analogs with peptide backbone instead of sugar backbone having high binding strength, specificity, and resistance to protease and nuclease degradation) [241, 242]. Antisense technology permits silencing a selected gene by binding to the mRNA produced by the gene and effectively inactivating the gene [243]. Endosomal escape following endocytotic entry is a critical step for these CPP conjugates to enter

cytosol/nucleus [244, 245]. The low efficiency in this step also forms the most important roadblock for intracellular delivery of CPP conjugates [246]. There have been demonstrations of improving the endosomal release by photodynamic treatment, Ca^{2+} treatment and chloroquine treatment [247, 248]. However, there has not been any quantitative assessment in the literature on the efficiency of the endosomal release of CPP conjugates and how much such process affects the eventual gene silencing result.

Antisense CPP-PNA conjugates have been used to inhibition of bacterial growth [249-252]. Delivery of naked PNAs into bacteria was typically limited by their outer-membrane lipopolysaccharide (LPS) layer [253, 254]. However, CPPs such as $(KFF)_3K$ could effectively breach the LPS layer and enable delivery of PNAs that target gene with sequence specificity for essential mRNA or rRNA sequences, inhibiting gene expression and thereby cell growth [252, 255, 256].

In spite of the success with using CPP-PNA conjugates on bacterial culture, there are additional layers of complexity when such strategy is applied to infected humans and animals. Pathogens like *Salmonella*, *Brucella, Listeria* and *Mycobacterium* have the unusual ability to not only survive in a host, but also subvert host immune responses. They can invade and live inside macrophages that are normally programmed to kill them [257]. Not surprisingly, $(KFF)_3K$-O-PNA conjugates showed much lower activity against bacteria encapsulated in membrane vesicles inside mammalian cells (e.g. macrophages) compared to direct use on bacterial culture, presumably due to the additional barriers formed by the plasma membrane and more importantly the endosomal membrane [258-260].

In this work, we delivered $(KFF)_3K$-O-PNA conjugates into murine macrophages J774.A cells that were infected with *Salmonella enterica* serovar Typhimurium LT2 (*S.*

Typhimurium LT2) [224]. Such infected systems are typical in animals and humans [261-267].

We found that the application of electroporation dramatically improved the bioavailability of (KFF)₃K-O-PNA conjugates inside the macrophages and produced substantially more inactivation of the encapsulated bacteria (up to a factor of 14) than the experiments without electroporation. Since electroporation is a technique that allows both *in vitro* and *in vivo* applications[268], we believe that the use of electroporation for CPP-PNA conjugates delivery will be an effective strategy for a wide range of applications. Our microfluidic electroporation device will permit functional screening of CPP-PNA systems based on their effectiveness for inhibition of intracellular bacteria.

7.2 Materials and Methods

Microfluidic channel fabrication

The *Salmonella* infected macrophages were electroporated in a microfluidic channel, which permitted counting of the viable macrophages before and after the treatment. The devices were made by standard soft lithography [269, 270]. Photomasks were designed by Freehand MX (Macromedia, San Francisco, CA) and printed on 4000dpi film. SU-8 2025 (MicroChem Corp., Newton, MA) was spincoated at 500 rpm for 10s and 1500rpm for 30 seconds in sequence on a 3-inch silicon wafer. The thickness of SU-8 was around 55µm. SU-8 layer on the silicon wafer was patterned under UV light and then developed.

Once the SU-8/silicon wafer master was fabricated, a two component silicone elastomer polydimethylsiloxane mixture (PDMS, GE RTV 615) with 10:1 ratio was poured on the master in a petri dish. The PDMS was vacuumed for 1h and cured in an oven for 1 h at 80℃. Then the PDMS structure was peeled off from the master and access holes

were punched. The PDMS structure and a pre-cleaned glass slide were then oxidized in a plasma cleaner (Harrick Plasma, Ithaca, NY) and brought into contact immediately after oxidization for irreversible bonding to form closed channels. After fabrication, the channels were immediately filled by 20 µl PBS buffer containing fibronectin (10µg/ml) in order to keep the channel hydrophilic and facilitate cell adherence.

CPP-PNA

(KFF)$_3$K-O-PNA (H-KFFKFFKFFK-o-cataagacggt-NH$_2$, PNA Bio Inc, CA) targeting an RNA polymerase sigma 70 (sigma D) factor gene (*rpo*D) demonstrated enhanced inhibition of *Salmonella* growth *in vitro* and insufficient inhibition in infected macrophages [258]. The PNA is conjugated with the CPP, (KFF)$_3$K, via O linker which enhances the solubility. (KFF)$_3$K-O-PNA is heated at 60℃ for 10 min before each use.

In vitro CPP-PNA Inhibition Assay

*S.*Typhimurium LT2 (ATCC 700720) was cultured in Tryptic Soy Broth (TSB) overnight at 37℃. The CPP-PNAs (0, 10, 15 and 20 µM) were incubated with 10^6 cfu/ml *Salmonella* in TSB for 8 h in a 96-well plate in triplicate. Samples were serially diluted and the colony forming units (CFUs) number of the surviving bacteria after each treatment were determined by subsequent plating onto Tryptic Soy Agar (TSA) plates.

Macrophage culture, infection and CPP-PNA treatment

A murine macrophage cell line (J774A.1) was cultured in DMEM medium supplemented with 10% Fetal Bovine Serum (FBS) and 1% Penicillin Streptomycin (PS). Macrophages were seeded at a density of 5×10^5 cells/ml in a 6-well plate in 2 ml. J774A.1

cells were infected with *S.* Typhimurium LT2 following the previously reported procedures [258, 271]. Briefly, overnight broth culture of *Salmonella* was centrifuged, resuspended in DMEM medium. Macrophages were infected at the multiplicity of infection (MOI) of 1:1 to 10:1 for 20 min. After 20 min infection, infected macrophages were washed with fresh DMEM containing 30µg/ml gentamicin to kill extracellular *Salmonella*. The medium was replaced with fresh DMEM containing (KFF)$_3$K-O-PNA and 10µg/ml gentamicin at 1 h post infection (h.p.i). Infected macrophages were incubated at 37℃, under 5% CO2. Macrophages were counted and lysed with PBS containing sodium deoxycholate (0.5%) at 4 h.p.i. The samples were serially diluted, and the numbers of *Salmonella* were determined by plating onto TSA.

Electroporation-assisted CPP-PNA treatment of infected macrophage cells

Electric pulses with a defined duration were generated by a relay (5501-05-1, Coto Technology, Kingstown, RI) controlled by a LabVIEW program via a data acquisition card (NI USB Digital I/O Device USB 6501) and a power supply (Bio-Rad, Hercules, CA) that provided a constant dc voltage (**Fig. 7.1**)[272]. The pulse duration was measured by an oscilloscope (B&K precision 2530, IL).

J774A.1 cells were infected with *Salmonella* as described above. Infected macrophages were harvested and suspended in fresh DMEM containing 30µg/ml gentamicin at a density of 10^7 cell/ml. This cell density ensured high seeding density in the microfluidic channel, at about 500-700 cells in each channel [273]. 20 µl macrophage cell suspension was loaded into each microfluidic channel that was pre-treated with fibronectin. The microfluidic chip was then incubated at 37℃, 5% CO$_2$ for 1 h in an

incubator to allow attachment of macrophages to the glass substrate. Image were taken to count macrophages in the microfluidic channel. The number of seeded macrophages in each channel was counted under a microscope. At 1 h.p.i., the infected cells were washed using 20µl of electroporation buffer (1mM $MgSO_4$, 8mM Na_2HPO_4, 2mM KH_2PO_4, 250mM surcose). Then the channel was filled with the electroporation buffer containing $(KFF)_3$K-O-PNA at a concentration of 30µM. Two platinum electrodes were inserted into the two reservoirs at the end of the channel (**Fig. 7.1**) and electric pulses of various durations and field intensities were applied to infected macrophages. After electroporation, the medium inside the channel was replaced with fresh DMEM containing 10µg/ml gentamicin and CPP-PNA. The reservoirs were sealed using microseal adhesive for PCR plates (Bio-Rad, Hercules, CA) to avoid evaporation. The chip was incubated at 37℃, 5% CO_2 for 4 h. Images were taken again to count the viable macrophages based on morphology. The number of bacterial cells was determined by lysing macrophages and conducting plate count.

7.3 Results and Discussion

Postulated mechanisms for (KFF)$_3$K-O-PNA conjugate delivery via electroporation and endocytosis.
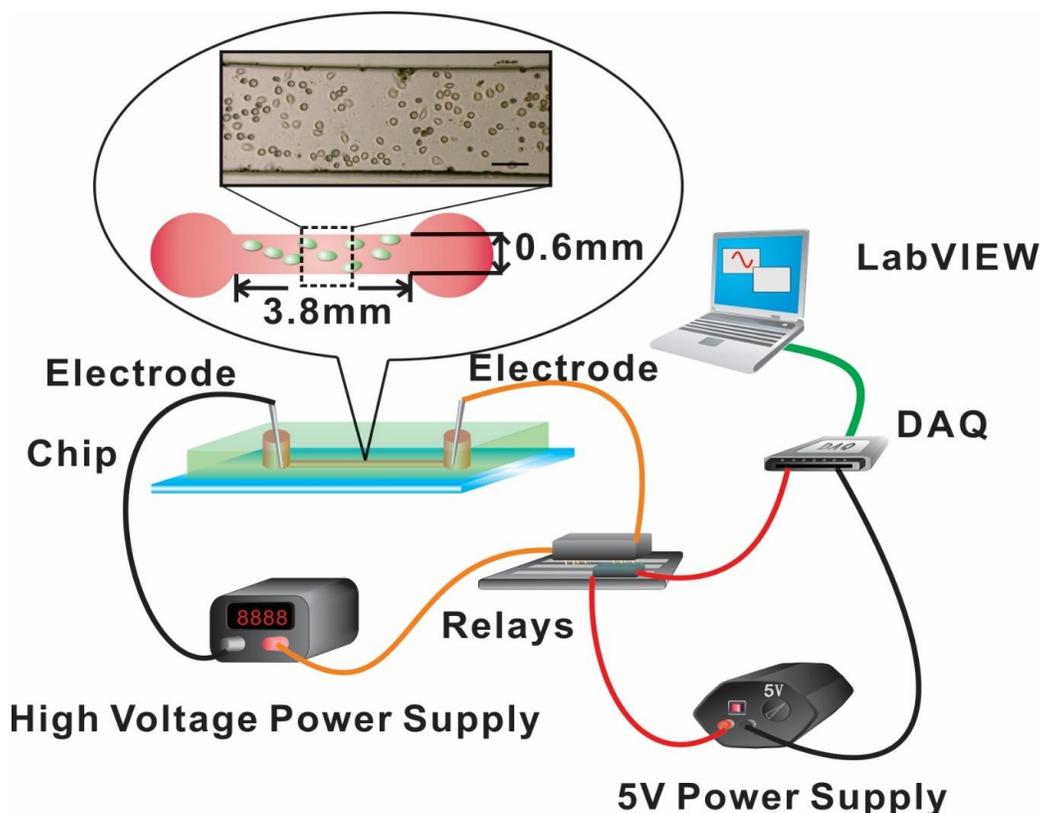
Fig. 7.1 Experimental setup for conducting electroporation-assisted CPP-PNA delivery. Electric pulse(s) was applied across a PDMS microfluidic channel via a pair of platinum (Pt) electrodes inserted in the reservoirs at both ends of the channel. The pulse duration was set by LabVIEW through DAQ and relays (supported by a 5V power supply). The field intensity of the pulse was determined by the voltage output from the high voltage power supply and the channel dimensions. The dimensions of each channel were 3.8mm×0.6mm×55µm. Each channel contained 500-700 macrophages. The scale bar in the phase contrast image is 200 µm.
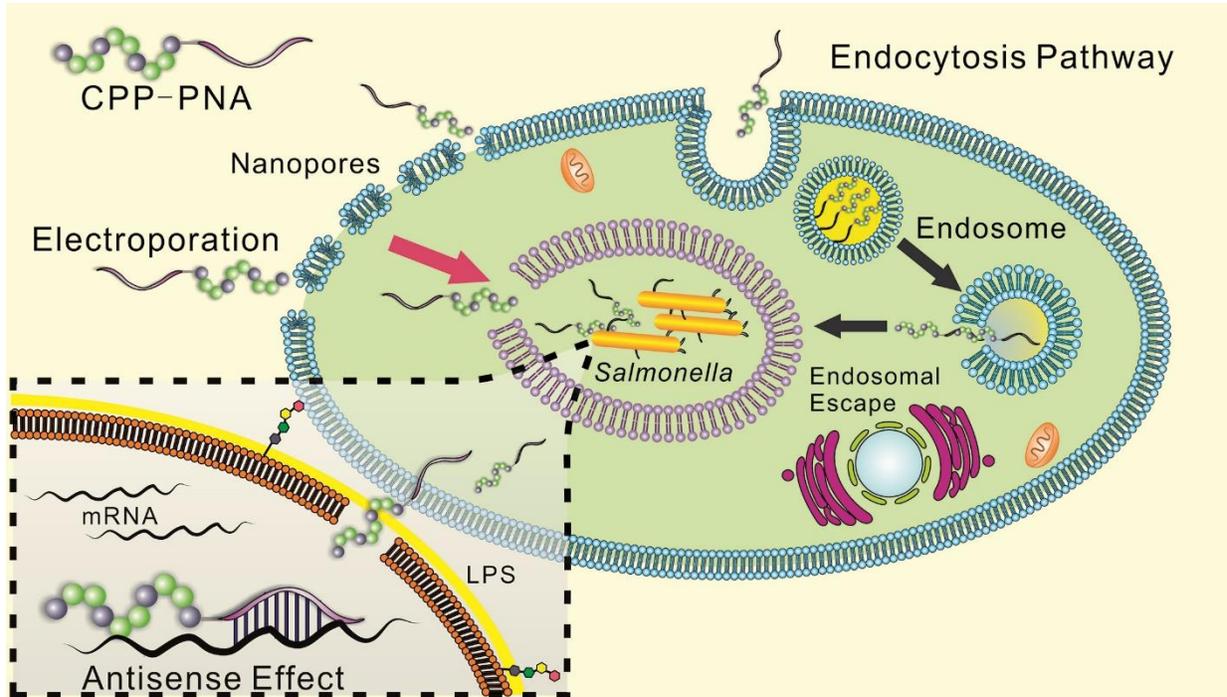
Fig. 7.2 Postulated mechanisms for (KFF)₃K-O-PNA conjugate delivery via electroporation and endocytosis into macrophages with lysozomal membrane encapsulated bacterial cells. (KFF)$_3$K-O-PNAs are encapsulated in endosomes after endocytosis into the macrophage cells. A fraction of the CPP-PNAs get released from endosomes inside the macrophage and then enter the encapsulated bacterial cells after penetrating the membrane and create antisense effects. In comparison, naked (KFF)$_3$K-O-PNAs enter the macrophage cell through electropores created by electroporation. They then move in the macrophage by diffusion and eventually enter the bacterial cells.

Electroporation (i.e. the exposure of cells to an external electric field with a higher-than-threshold intensity for a short duration) breaches cell membranes by creating nanoscale pores and destabilizing the membrane structures [13, 269, 274, 275]. Macromolecules in the solution such as nucleic acids or proteins become bound to the

breached membrane during electroporation process and they move into the cytosol or nucleus driven by a combination of electrophoresis and diffusion [13, 274-277]. Electroporation process does not involve formation of endosomes and endocytosis as a part of delivery. In comparison, although the mechanism for (KFF)$_3$K-O-PNA conjugates delivery is not completely understood, such processes are believed to involve endocytosis which involves encapsulation of the molecules in vesicles [248, 260]. The release of the conjugates from the endosomes is critical for the bioavailability and antisense efficiency for bacterial inhibition. The comparison of the two mechanisms is shown in **Fig. 7.2**.

Fig. 7.3 shows the significant difference in the antisense activity of (KFF)$_3$K-O-PNA toward *Salmonella* culture in *vitro* and *Salmonella* after infection of macrophages. To prove CPP-PNA can enter and inhibit *Salmonella* effectively, off-chip inhibition assay was applied. (**Fig. 7.3**) *In vitro Salmonella* culture and infected macrophages were treated with (KFF)$_3$K-O-PNA targeting an RNA polymerase sigma 70 (sigma D) factor gene *rpoD* which was essential for *Salmonella* replication [258]. For *in vitro* bacterium, various concentrations of the conjugate were applied for 8 h in 96-well plates. The CPP-PNA conjugate had efficient antisense effect that was concentration-dependent. Compared to the control (without application of the antisense reagent), the bacterial concentration was lower by 7 log reduction at 20 µM CPP-PNA concentration. In comparison, we applied the same antisense reagent to *Salmonella* infected J774A.1 macrophages and then examined the change in the bacterium/macrophage ratio at 4 h.p.i. (i.e. the number of *Salmonella* cells per macrophage cell). **Fig. 7.3** shows that the increase in the CPP-PNA concentration from 0 to 45 µM decreased the *Salmonella/macrophage* ratio from 3.3 to 1.5. The increase in the concentration beyond 30 µM did not appear to decrease the ratio

further. Our results confirm that the encapsulation of *Salmonella* with endosomal membrane inside macrophages greatly decreases the bioavailability of CPP-PNA toward bacterial inhibition.

(a)

(b)
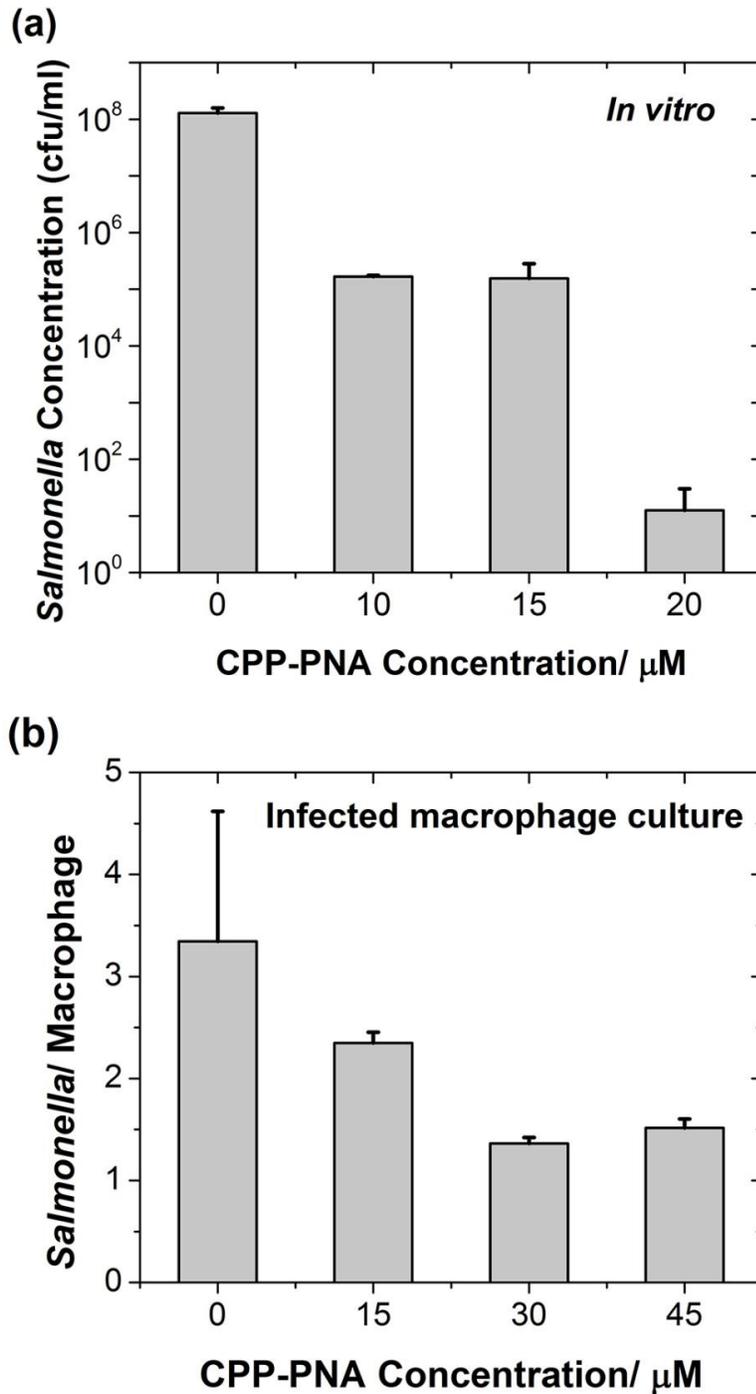
Fig. 7.3 Antisense effect of (KFF)₃K-O-PNAs in vitro (i.e. treating Salmonella culture directly) (a) and toward infected macrophages (b). (KFF)₃K-O-PNA designed to target rpoD gene was used in all experiments. The macrophage cells were infected with multiplicity of infection (moi) of 4:1.

(a)

(b)
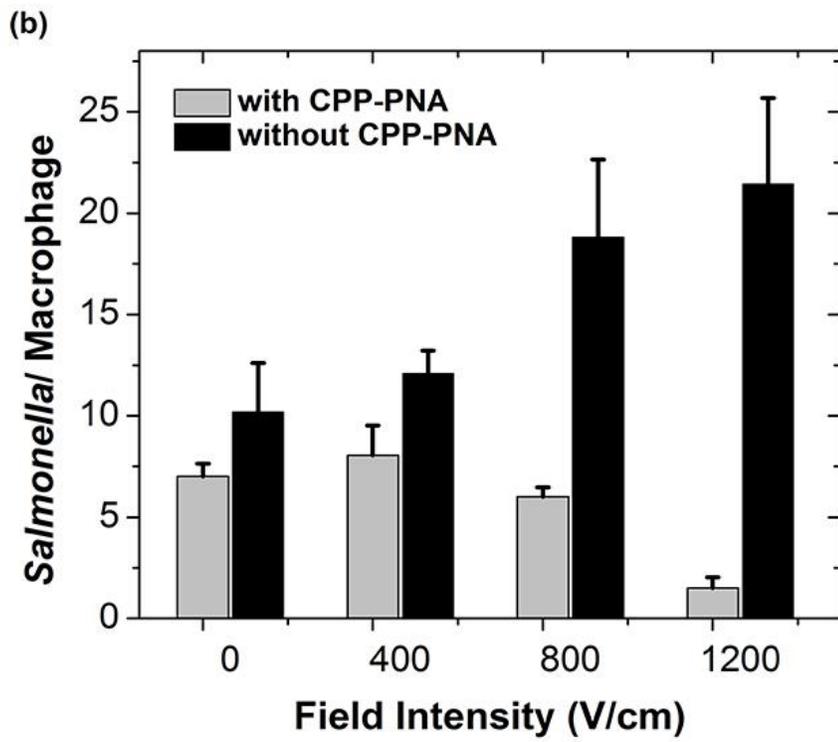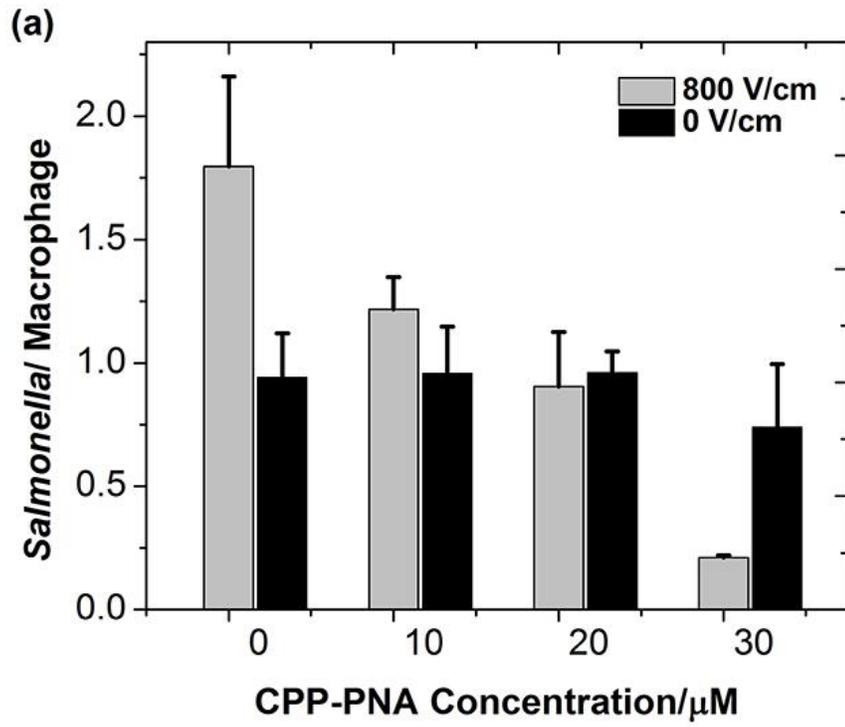
Fig. 7.4 The effect of electroporation on the antisense activity of (KFF)₃K-O-PNA.

Roughly 1,000 macrophages were evaluated for each data point. (a) The average

number of *Salmonella* per macrophage under various (KFF)₃K-O-PNA concentrations

before and after electroporation (with a single pulse of 800 V/cm and 5 ms). MOI of

1:1 was used in the infection. (b) The average number of *Salmonella* per macrophage

under electroporation of various intensities (with a single pulse of 5ms) with and

without (KFF)₃K-O-PNA (at a concentration of 30 µM). MOI of 10:1 was used in the

infection.

Electroporation significantly improves the bioavailability of CPP-PNA inside macrophages.

We applied single-pulse electroporation of various field intensities to macrophages

infected with various amounts of *Salmonella,* while having a CPP-PNA concentration of

30 µM. **Fig. 7.4** shows that electroporation significantly improves the bioavailability of

CPP-PNA inside macrophages. **Fig. 7.4a** shows that under electroporation (one single

pulse of 800V/cm and 5 ms), the average number of *Salmonella* per macrophage cell

decreased consistently and drastically (from 1.8 at 0µM to 0.2 at 30µM PNA) while CPP-

PNA concentration increased. The data indicates that the CPP-PNA delivery was not

saturated at the concentrations 20-30 µM under the electroporation condition. In

comparison, when there was no electroporation applied, there was only a modest

decrease (~21%) in the *Salmonella*/macrophage ratio when the CPP-PNA concentration

increased up to 30 µM. **Fig. 7.4b** shows that the increase in the field intensity of the pulse

(from 0 to 1200 V/cm) also created decrease in the *Salmonella*/macrophage ratio when

CPP-PNA is delivered, due to increased delivery. However, stronger field intensity also

led to death of macrophages (**Fig. 7.5**) thus increased the *Salmonella*/macrophage ratio

when CPP-PNA was not present. The comparison between **Fig. 7.4b** and **Fig. 7.5** reveals that *Salmonella* was not killed by the electric field at 0-800 V/cm because the death rates of the macrophages (shown in **Fig. 7.5)** solely contributed to the increase in the *Salmonella* /macrophage ratio (shown in **Fig. 7.5b,** without CPP-PNA). However, roughly 23% of total *Salmonella* died due to electroporation under 1200 V/cm pulse. This threshold for electric lysis of *Salmonella* is similar to the one we previously reported for *E. coli* [278].

Decrease in the average Salmonella per macrophage under various infection and electroporation conditions.

In **Fig. 7.6**, we examine the decrease in the average *Salmonella* per macrophage due to the antisense inhibition (i.e. the difference in the *Salmonella*/Macrophage ratio between the cases of with and without CPP-PNA) under various infection and electroporation conditions. Electroporation is effective for the antisense inhibition in macrophage infected by various amounts of *Salmonella*. The antisense inhibition increased with higher field intensity used in the electroporation. Furthermore, the more heavily infected macrophages experienced more significant reduction in the average bacterial cell number per macrophage cell.
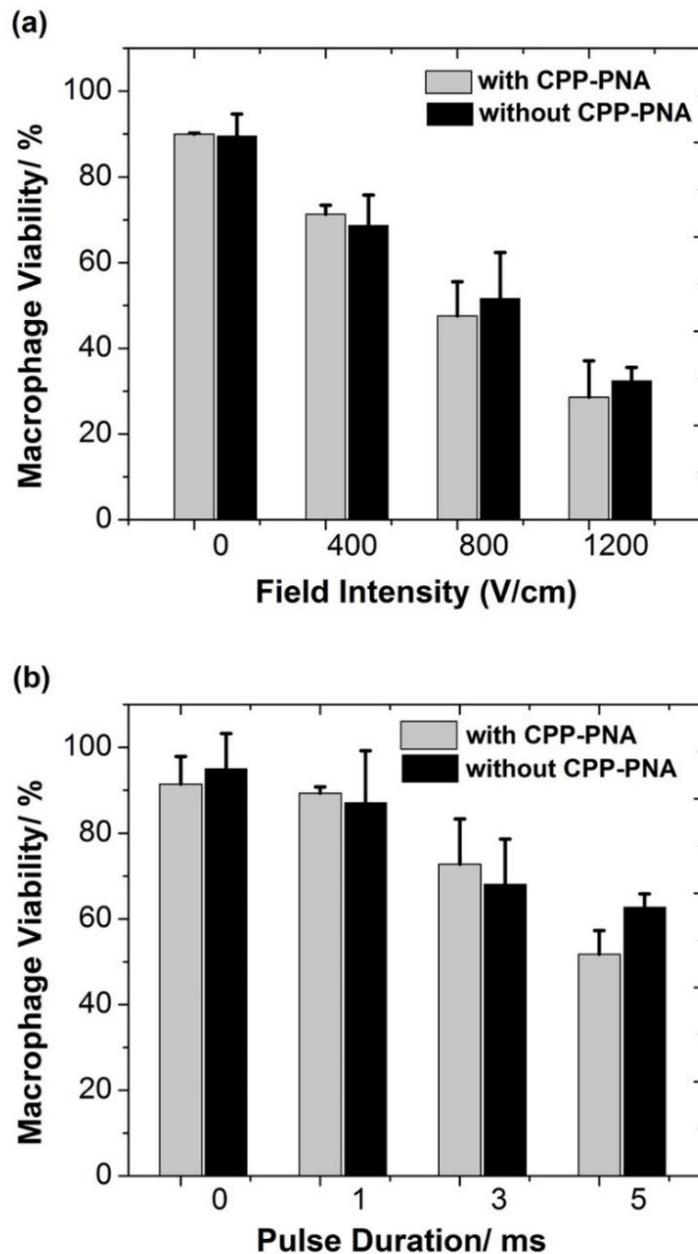
Fig. 7.5 Macrophage viability under various electroporation conditions. The presence of CPP-PNA (30μM) did not affect the macrophage viability. Macrophages were not infected. (a) Macrophage viability under various field intensities when a single pulse with 5 ms duration was applied. (b) Macrophage viability under various pulse durations. A single pulse with 800V/cm field intensity was applied in all tests.

7.4 Conclusion

To summarize, we found that electroporation dramatically increases the bioavailability of (KFF)$_3$K-O-PNA for inhibiting intracellular *Salmonella* inside macrophages. More intense electroporation (i.e. higher field intensity and longer duration) leads to more effective antisense activity but also higher macrophage death rate. We postulate that the improvement in the antisense effect of CPP-PNAs by electroporation was due to elimination of endocytic delivery that involves cargo encapsulation inside endosomes. We believe that electroporation can be widely applied to improve CPP-PNA delivery into cells, especially in complex cellular systems involving multiple barriers. Microfluidic electroporation permits the screening of these antisense reagents on a convenient platform.
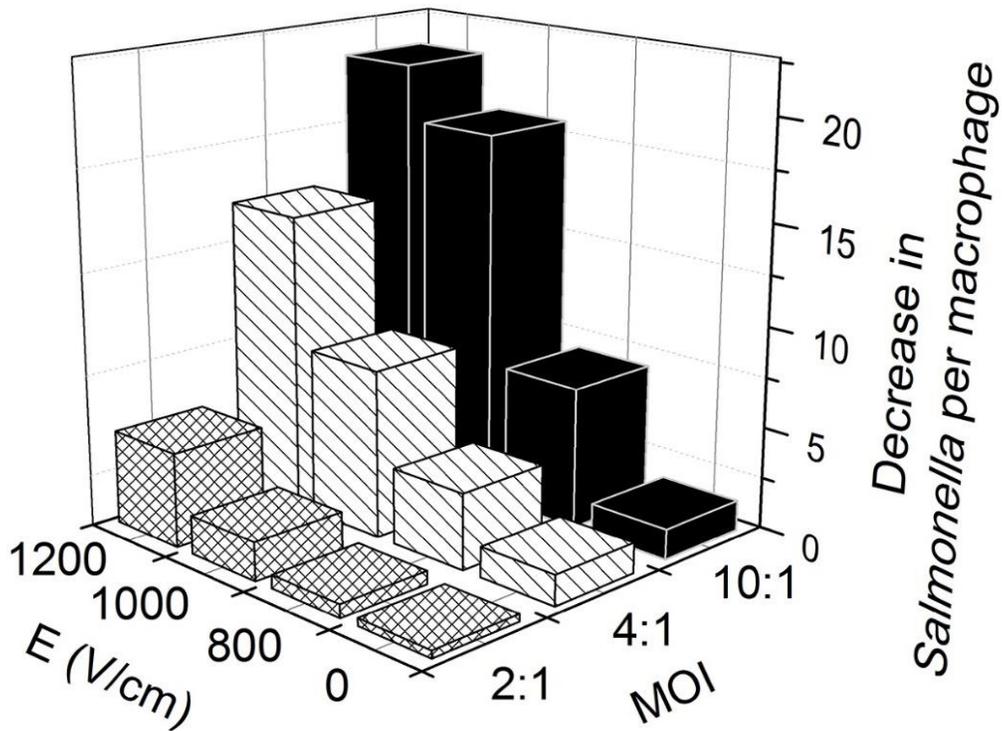
Fig. 7.6 The antisense effect (measured by the decrease in the average *Salmonella* **number per macrophage due to the addition of (KFF)₃K-O-PNA)** **under various electroporation and infection conditions.** Macrophages with various multiplicity of infection (2:1 to 10:1) were treated with a single 5 ms electric pulse and 30 μM (KFF)₃K-O-PNA. The decrease in the average *Salmonella* per macrophage was calculated by examining the difference in the ratio *Salmonella*/macrophage between the test with (KFF)₃K-O-PNA application and the one without its application.

CHAPTER 8 – SUMMARY AND OUTLOOK

Microfluidics offers new opportunities and insights to the genomics and epigenomics. In this thesis, I described my projects in developing microfluidic assays genomic (PCR) and epigenomic assays (MID-RRBS and SurfaceChIP-seq). Numerous applications of microfluidics in the field of epigenomics have not been explored yet. The future work will potentially focus on three aspects.

Firstly, epigenomic assays for RNA-protein interaction, RNA methylation, DNA methylation vatients (5-hmC, 5-fc, and 5-caC), and transcription factor binding, are not adapted to microfluidic platform. There is no doubt that these assays will be adapted to microfluidic devices in the near future. Another challenge of current assays are the sensitivity. Unlike genomic and transcriptomic assays, most epigenomic assays require lot of input materials. Microfluidics show great potential in manipulating low amount of cells, even single cells. Utilizing microfluidics, a few epigenomic assays achieved single cell sensitivity, such as ChIP-seq and ATAC-seq. It highly likely that we will see more single-cell assays. Lastly, one of goal of epigenomics is to find out the connection between epigenomic signal and phenotype. To build up the intact landscape of a single cell, it is necessary to integrate multiple assays. Several efforts integrate DNA methylation/gene expression or DNA methylation/histone modification. Although challenging, more assays, potentially single-cell assays, will be integrated which would reveal the genuine molecular mechanisms.

REFERENCES

1. Cusanovich, D.A. et al. Epigenetics. Multiplex single-cell profiling of chromatin accessibility by combinatorial cellular indexing. *Science* **348**, 910-914 (2015).

2. Hough, Shelley R. et al. Single-Cell Gene Expression Profiles Define Self-Renewing, Pluripotent, and Lineage Primed States of Human Pluripotent Stem Cells. *Stem Cell Reports* **2**, 881-895.

3. Klein, A.M. et al. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* **161**, 1187-1201 (2015).

4. Kumar, R.M. et al. Deconstructing transcriptional heterogeneity in pluripotent stem cells. *Nature* **516**, 56-61 (2014).

5. Patel, A.P. et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* **344**, 1396-1401 (2014).

6. Kalisky, T. & Quake, S.R. Single-cell genomics. *Nat. Methods* **8**, 311-314 (2011).

7. Saliba, A.-E., Westermann, A.J., Gorski, S.A. & Vogel, J. Single-cell RNA-seq: advances and future challenges. *Nucleic Acids Res.* **42**, 8845-8860 (2014).

8. Nawy, T. Single-cell sequencing. *Nat. Methods* **11**, 18-18 (2014).

9. Wu, A.R. et al. Quantitative assessment of single-cell RNA-sequencing methods. *Nat. Methods* **11**, 41-46 (2014).

10. Kalisky, T., Blainey, P. & Quake, S.R. Genomic analysis at the single-cell level. *Annu. Rev. Genomics Hum. Genet.* **45**, 431-445 (2011).

11. Gawad, C., Koh, W. & Quake, S.R. Single-cell genome sequencing: current state of the science. *Nat. Rev. Genet.* **17**, 175-188 (2016).

12.   Sun, C. et al. Electroporation-delivered fluorescent protein biosensors for probing molecular activities in cells without genetic encoding. *Chem. Comm.* **50**, 11536-11539 (2014).

13.   Geng, T. & Lu, C. Microfluidic electroporation for cellular analysis and delivery. *Lab Chip* **13**, 3803-3821 (2013).

14.   Sun, C. et al. Paramagnetic Structures within a Microfluidic Channel for Enhanced Immunomagnetic Isolation and Surface Patterning of Cells. *Sci. Rep.* **6**, 29407 (2016).

15.   Loufakis, D.N., Cao, Z., Ma, S., Mittelman, D. & Lu, C. Focusing of mammalian cells under an ultrahigh pH gradient created by unidirectional electropulsation in a confined microchamber. *Chem. Sci.* **5**, 3331-3337 (2014).

16.   Ma, S., Tang, Y., Liu, J. & Wu, J. Visible paper chip immunoassay for rapid determination of bacteria in water distribution system. *Talanta* **120**, 135-140 (2014).

17.   Ma, S., Murphy, T.W. & Lu, C. Microfluidics for genome-wide studies involving next generation sequencing. *Biomicrofluidics* **11**, 021501 (2017).

18.   Tan, J. et al. Capture, enrichment, and mass spectrometric detection of low-molecular-weight biomarkers with nanoporous silicon microparticles. *Adv. Healthc. Mater.* **1**, 742-750 (2012).

19.   Sun, C. et al. Immunomagnetic separation of tumor initiating cells by screening two surface markers. *Sci. Rep.* **7**, 40632 (2017).

20.   Handy, D.E., Castro, R. & Loscalzo, J. Epigenetic Modifications: Basic Mechanisms and Role in Cardiovascular Disease. *Circulation* **123**, 2145-2156 (2011).

21.   Lister, R. et al. Global epigenomic reconfiguration during mammalian brain development. *Science* **341**, 1237905 (2013).

22.   Sharma, S., Kelly, T.K. & Jones, P.A. Epigenetics in cancer. *Carcinogenesis* **31**, 27-36 (2010).

23.   Oh, H.J. et al. DNA-Enrichment Microfluidic Chip for Chromatin Immunoprecipitation. *Anal. Chem.* **81**, 2832-2839 (2009).

24.   Wu, A.R. et al. Automated microfluidic chromatin immunoprecipitation from 2,000 cells. *Lab Chip* **9**, 1365-1370 (2009).

25.   Wu, A.R. et al. High throughput automated chromatin immunoprecipitation as a platform for drug screening and antibody validation. *Lab Chip* **12**, 2190-2198 (2012).

26.   Geng, T. et al. Histone modification analysis by chromatin immunoprecipitation from a low number of cells on a microfluidic platform. *Lab Chip* **11**, 2842-2848 (2011).

27.   Cao, Z. & Lu, C. A Microfluidic Device with Integrated Sonication and Immunoprecipitation for Sensitive Epigenetic Assays. *Anal. Chem.* **88**, 1965-1972 (2016).

28.   Shen, J. et al. H3K4me3 epigenomic landscape derived from ChIP-Seq of 1 000 mouse early embryonic cells. *Cell Res.* **25**, 143-147 (2015).

29.   Cao, Z., Chen, C., He, B., Tan, K. & Lu, C. A microfluidic device for epigenomic profiling using 100 cells. *Nat. Methods* **12**, 959-962 (2015).

30.   Buenrostro, J.D. et al. Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature* **523**, 486-490 (2015).

31.   Rotem, A. et al. Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. *Nat. Biotechnol.* **33**, 1165-1172 (2015).

32.   Khavari, D.A., Sen, G.L. & Rinn, J.L. DNA methylation and epigenetic control of cellular differentiation. *Cell Cycle* **9**, 3880-3883 (2010).

33.   Sheaffer, K.L. et al. DNA methylation is required for the control of stem cell differentiation in the small intestine. *Genes Dev.* **28**, 652-664 (2014).

34.   Brinkman, A.B. et al. Sequential ChIP-bisulfite sequencing enables direct genome-scale investigation of chromatin and DNA methylation cross-talk. *Genome Res.* **22**, 1128-1138 (2012).

35.   Statham, A.L. et al. Bisulfite sequencing of chromatin immunoprecipitated DNA (BisChIP-seq) directly informs methylation status of histone-modified DNA. *Genome Res.* **22**, 1120-1127 (2012).

36.   O'Neill, L.P., VerMilyea, M.D. & Turner, B.M. Epigenetic characterization of the early embryo with a chromatin immunoprecipitation protocol applicable to small cell populations. *Nat. Genet.* **38**, 835-841 (2006).

37.   Adli, M. & Bernstein, B.E. Whole-genome chromatin profiling from limited numbers of cells using nano-ChIP-seq. *Nat. Protoc.* **6**, 1656-1668 (2011).

38.   Adli, M., Zhu, J. & Bernstein, B.E. Genome-wide chromatin maps derived from limited numbers of hematopoietic progenitors. *Nat. Methods* **7**, 615-618 (2010).

39.   Shankaranarayanan, P. et al. Single-tube linear DNA amplification (LinDA) for robust ChIP-seq. *Nat. Methods* **8**, 565-567 (2011).

40.   Lara-Astiaso, D. et al. Immunogenetics. Chromatin state dynamics during blood formation. *Science* **345**, 943-949 (2014).

41.   Brind'Amour, J. et al. An ultra-low-input native ChIP-seq protocol for genome-wide profiling of rare cell populations. *Nat. Commun.* **6** (2015).

42.   Ng, J.H. et al. In vivo epigenomic profiling of germ cells reveals germ cell molecular signatures. *Dev. Cell* **24**, 324-333 (2013).

43.   Zheng, X. et al. Low-Cell-Number Epigenome Profiling Aids the Study of Lens Aging and Hematopoiesis. *Cell Rep.* **13**, 1505-1518 (2015).

44.   Zwart, W. et al. A carrier-assisted ChIP-seq method for estrogen receptor-chromatin interactions from breast cancer core needle biopsy samples. *BMC Genomics* **14**, 232 (2013).

45.   Jung, Y.L. et al. Impact of sequencing depth in ChIP-seq experiments. *Nucleic Acids Res.* **42**, e74 (2014).

46.   Grada, A. & Weinbrecht, K. Next-Generation Sequencing: Methodology and Application. *J. Invest. Dermatol.* **133**, e11 (2013).

47.   Meissner, A. et al. Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Res.* **33**, 5868-5877 (2005).

48.   Zilberman, D., Gehring, M., Tran, R.K., Ballinger, T. & Henikoff, S. Genome-wide analysis of Arabidopsis thaliana DNA methylation uncovers an interdependence between methylation and transcription. *Nat. Genet.* **39**, 61-69 (2007).

49.   Lister, R. et al. Highly integrated single-base resolution maps of the epigenome in Arabidopsis. *Cell* **133**, 523-536 (2008).

50.   Guo, H. et al. Single-cell methylome landscapes of mouse embryonic stem cells and early embryos analyzed using reduced representation bisulfite sequencing. *Genome Res.* **23**, 2126-2135 (2013).

51.   Smallwood, S.A. et al. Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. *Nat. Methods* **11**, 817-820 (2014).

52.   Hou, Y. et al. Single-cell triple omics sequencing reveals genetic, epigenetic, and transcriptomic heterogeneity in hepatocellular carcinomas. *Cell Res.* **26**, 304-319 (2016).

53.   Clark, S.J., Lee, H.J., Smallwood, S.A., Kelsey, G. & Reik, W. Single-cell epigenomics: powerful new methods for understanding gene regulation and cell identity. *Genome Biol.* **17**, 1-10 (2016).

54.   Stark, A., Shin, D.J., Pisanic, T., 2nd, Hsieh, K. & Wang, T.H. A parallelized microfluidic DNA bisulfite conversion module for streamlined methylation analysis. *Biomed. Microdevices* **18**, 5 (2016).

55.   Yoon, J., Park, M.K., Lee, T.Y., Yoon, Y.-J. & Shin, Y. LoMA-B: a simple and versatile lab-on-a-chip system based on single-channel bisulfite conversion for DNA methylation analysis. *Lab Chip* **15**, 3530-3539 (2015).

56.   Komori, H.K. et al. Application of microdroplet PCR for large-scale targeted bisulfite sequencing. *Genome Res.* **21**, 1738-1745 (2011).

57.   Herrmann, A. et al. Pipeline for large-scale microdroplet bisulfite PCR-based sequencing allows the tracking of hepitype evolution in tumors. *PLoS One* **6**, e21332 (2011).

58.   Zhang, Y. et al. DNA methylation analysis on a droplet-in-oil PCR array. *Lab Chip* **9**, 1059-1064 (2009).

59.   Weisenberger, D.J. et al. DNA methylation analysis by digital bisulfite genomic sequencing and digital MethyLight. *Nucleic Acids Res.* **36**, 4689-4698 (2008).

60.   Gilmour, D.S. & Lis, J.T. In vivo interactions of RNA polymerase II with genes of Drosophila melanogaster. *Mol. Cell. Biol.* **5**, 2009-2018 (1985).

61.   Matsuoka, T., Choul Kim, B., Moraes, C., Han, M. & Takayama, S. Micro- and nanofluidic technologies for epigenetic profiling. *Biomicrofluidics* **7**, 041301 (2013).

62.   Zhao, M.T., Whyte, J.J., Hopkins, G.M., Kirk, M.D. & Prather, R.S. Methylated DNA immunoprecipitation and high-throughput sequencing (MeDIP-seq) using low amounts of genomic DNA. *Cell. Reprogram.* **16**, 175-184 (2014).

63.   Down, T.A. et al. A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis. *Nat. Biotechnol.* **26**, 779-785 (2008).

64.   Aberg, K.A. et al. MBD-seq as a cost-effective approach for methylome-wide association studies: demonstration in 1500 case--control samples. *Epigenomics* **4**, 605-621 (2012).

65.   Lee, T.Y., Shin, Y. & Park, M.K. A simple, low-cost, and rapid device for a DNA methylation-specific amplification/detection system using a flexible plastic and silicon complex. *Lab Chip* **14**, 4220-4229 (2014).

66.   De, A., Sparreboom, W., van den Berg, A. & Carlen, E.T. Rapid microfluidic solid-phase extraction system for hyper-methylated DNA enrichment and epigenetic analysis. *Biomicrofluidics* **8**, 054119 (2014).

67.   Kurita, R. & Niwa, O. Microfluidic platforms for DNA methylation analysis. *Lab Chip*
        16, 3631-3644 (2016).

68.   Nelson, C.S. et al. Microfluidic affinity and ChIP-seq analyses converge on a
        conserved FOXP2-binding motif in chimp and human, which enables the detection
        of evolutionarily novel targets. *Nucleic Acids Res.* **41**, 5991-6004 (2013).

69.   Maerkl, S.J. & Quake, S.R. A Systems Approach to Measuring the Binding Energy
        Landscapes of Transcription Factors. *Science* **315**, 233-237 (2007).

70.   Fordyce, P.M. et al. De novo identification and biophysical characterization of
        transcription-factor binding sites with microfluidic affinity analysis. *Nat. Biotechnol.*
        28, 970-975 (2010).

71.   Glick, Y. et al. Integrated microfluidic approach for quantitative high-throughput
        measurements of transcription factor binding affinities. *Nucleic Acids Res.* **44**, e51
        (2015).

72.   Rockel, S., Geertz, M. & Maerkl, S.J. MITOMI: a microfluidic platform for in vitro
        characterization of transcription factor-DNA interaction. *Methods. Mol. Biol.* **786**,
        97-114 (2012).

73.   Warren, L., Bryder, D., Weissman, I.L. & Quake, S.R. Transcription factor profiling
        in individual hematopoietic progenitors by digital RT-PCR. *Proc. Natl. Acad. Sci.*
        U.S.A. **103**, 17807-17812 (2006).

74.   Meyer, C.A. & Liu, X.S. Identifying and mitigating bias in next-generation
        sequencing methods for chromatin biology. *Nat. Rev. Genet.* **15**, 709-721 (2014).

75.   Hakim, O. & Misteli, T. SnapShot: Chromosome confirmation capture. *Cell* **148**,
        1068.e1061-1062 (2012).

76.   Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y. & Greenleaf, W.J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213-1218 (2013).

77.   Boyle, A.P. et al. High-resolution mapping and characterization of open chromatin across the genome. *Cell* **132**, 311-322 (2008).

78.   Barski, A. et al. High-resolution profiling of histone methylations in the human genome. *Cell* **129**, 823-837 (2007).

79.   Giresi, P.G., Kim, J., McDaniell, R.M., Iyer, V.R. & Lieb, J.D. FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. *Genome Res.* **17**, 877-885 (2007).

80.   Nagano, T. et al. Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* **502**, 59-64 (2013).

81.   Jin, W. et al. Genome-wide Detection of DNase I Hypersensitive Sites in Single Cells and FFPE Samples. *Nature* **528**, 142-146 (2015).

82.   Small, E.C., Xi, L., Wang, J.-P., Widom, J. & Licht, J.D. Single-cell nucleosome mapping reveals the molecular basis of gene expression heterogeneity. *Proc. Natl. Acad. Sci. U.S.A.* **111**, e2462-e2471 (2014).

83.   Geng, T., Bao, N., Sriranganathanw, N., Li, L. & Lu, C. Genomic DNA Extraction from Cells by Electroporation on an Integrated Microfluidic Platform. *Anal. Chem.* **84**, 9632-9639 (2012).

84.   Hall, E.W., Kim, S., Appadoo, V. & Zare, R.N. Lysis of a single cyanobacterium for whole genome amplification. *Micromachines* **4**, 321-332 (2013).

85. Geng, T., Bao, N., Gall, O.Z. & Lu, C. Modulating DNA adsorption on silica beads using an electrical switch. *Chem. Comm.*, 800-802 (2009).

86. Ma, S., Bryson, B.D., Sun, C., Fortune, S.M. & Lu, C. RNA Extraction from a Mycobacterium under Ultrahigh Electric Field Intensity in a Microfluidic Device. *Anal. Chem.* **88**, 5053-5057 (2016).

87. Liu, C.J. et al. Magnetic-bead-based microfluidic system for ribonucleic acid extraction and reverse transcription processes. *Biomed. Microdevices* **11**, 339-350 (2009).

88. Irimia, D. et al. Genome-wide transcriptome analysis of 150 cell samples. *Integr. Biol.* **1**, 99-107 (2009).

89. Jebrail, M.J. et al. World-to-Digital-Microfluidic Interface Enabling Extraction and Purification of RNA from Human Whole Blood. *Anal. Chem.* **86**, 3856-3862 (2014).

90. Han, N., Shin, J.H. & Han, K.-H. An on-chip RT-PCR microfluidic device, that integrates mRNA extraction, cDNA synthesis, and gene amplification. *RSC Adv.* **4**, 9160-9165 (2014).

91. Bao, N., Wang, J. & Lu, C. Recent advances in electric analysis of cells in microfluidic systems. *Anal. Bioanal. Chem.* **391**, 933-942 (2008).

92. Wu, J., Kodzius, R., Cao, W. & Wen, W. Extraction, amplification and detection of DNA in microfluidic chip-based assays. *Microchim. Acta* **181**, 1611-1631 (2014).

93. Cui, F., Rhee, M., Singh, A. & Tripathi, A. Microfluidic Sample Preparation for Medical Diagnostics. *Annu. Rev. Biomed. Eng.* **17**, 267-286 (2015).

94. Horsman, K.M., Bienvenue, J.M., Blasier, K.R. & Landers, J.P. Forensic DNA analysis on microfluidic devices: a review. *J. Forensic Sci.* **52**, 784-799 (2007).

95.   Sista, R. et al. Development of a digital microfluidic platform for point of care testing. *Lab Chip* **8**, 2091-2104 (2008).

96.   Choi, K., Ng, A.H., Fobel, R. & Wheeler, A.R. Digital microfluidics. *Annu. Rev. Anal. Chem.* **5**, 413-440 (2012).

97.   Tseng, Q., Lomonosov, A.M., Furlong, E.E. & Merten, C.A. Fragmentation of DNA in a sub-microliter microfluidic sonication device. *Lab Chip* **12**, 4677-4682 (2012).

98.   Miyazaki, K. Random DNA fragmentation with endonuclease V: application to DNA shuffling. *Nucleic Acids Res.* **30**, e139 (2002).

99.   Fan, H.C., Wang, J., Potanina, A. & Quake, S.R. Whole-genome molecular haplotyping of single cells. *Nat. Biotechnol.* **29**, 51-57 (2011).

100.   Kumar, R. et al. A high-throughput method for Illumina RNA-Seq library preparation. *Front. Plant Sci.* **3**, 202 (2012).

101.   Knierim, E., Lucke, B., Schwarz, J.M., Schuelke, M. & Seelow, D. Systematic Comparison of Three Methods for Fragmentation of Long-Range PCR Products for Next Generation Sequencing. *PLoS One* **6**, e28240 (2011).

102.   Anderson, R.C., Su, X., Bogdan, G.J. & Fenton, J. A miniature integrated device for automated multistep genetic assays. *Nucleic Acids Res.* **28**, E60 (2000).

103.   Jong Wook, H., Yan, C., Anderson, W.F. & Stephen, R.Q. Molecular biology on a microfluidic chip. *J. Phys. Condens. Matter* **18**, S691 (2006).

104.   Lin, H.-C., Liu, Y.-J. & Yao, D.-J. Core—Shell Droplets for Parallel DNA Ligation of an Ultra-Micro Volume Using an EWOD Microfluidic System. *J. Lab. Autom.* **15**, 210-215 (2010).

105. Ko, Y.-J., Maeng, J.-H., Ahn, Y. & Hwang, S.Y. DNA ligation using a disposable microfluidic device combined with a micromixer and microchannel reactor. *Sensor. Actuat. B-Chem.* **157**, 735-741 (2011).

106. Kim, H. et al. Automated Digital Microfluidic Sample Preparation for Next-Generation DNA Sequencing. *J. Lab. Autom.* **16**, 405-414 (2011).

107. Kim, H. et al. A microfluidic DNA library preparation platform for next-generation sequencing. *PLoS One* **8**, e68988 (2013).

108. Tan, S.J. et al. A microfluidic device for preparing next generation DNA sequencing libraries and for automating other laboratory protocols that require one or more column chromatography steps. *PLoS One* **8**, e64084 (2013).

109. White, R.A., Blainey, P.C., Fan, H.C. & Quake, S.R. Digital PCR provides sensitive and absolute calibration for high throughput sequencing. *BMC Genomics* **10**, 1-12 (2009).

110. Shen, F., Du, W., Kreutz, J.E., Fok, A. & Ismagilov, R.F. Digital PCR on a SlipChip. *Lab Chip* **10**, 2666-2672 (2010).

111. Hindson, B.J. et al. High-Throughput Droplet Digital PCR System for Absolute Quantitation of DNA Copy Number. *Anal. Chem.* **83**, 8604-8610 (2011).

112. Whale, A.S. et al. Comparison of microfluidic digital PCR and conventional quantitative PCR for measuring copy number variation. *Nucleic Acids Res.* **40**, e82 (2012).

113. Morley, A.A. Digital PCR: A brief history. *Biomol. Detect. Quantif.* **1**, 1-2 (2014).

114. Pohl, G. & Shih Ie, M. Principle and applications of digital PCR. *Expert Rev. Mol. Diagn.* **4**, 41-47 (2004).

115.   Dong, L. et al. Comparison of four digital PCR platforms for accurate quantification of DNA copy number of a certified plasmid DNA reference material. *Sci. Rep.* **5**, 13174 (2015).

116.   Thaitrong, N. et al. Quality control of next-generation sequencing library through an integrative digital microfluidic platform. *Electrophoresis* **33**, 3506-3513 (2012).

117.   Ottesen, E.A., Hong, J.W., Quake, S.R. & Leadbetter, J.R. Microfluidic Digital PCR Enables Multigene Analysis of Individual Environmental Bacteria. *Science* **314**, 1464-1467 (2006).

118.   Marcus, J.S., Anderson, W.F. & Quake, S.R. Microfluidic Single-Cell mRNA Isolation and Analysis. *Anal. Chem.* **78**, 3084-3089 (2006).

119.   Easley, C.J. et al. A fully integrated microfluidic genetic analysis system with sample-in–answer-out capability. *P. Natl. Acad. Sci. USA* **103**, 19272-19277 (2006).

120.   Muddu, R., Hassan, Y.A. & Ugaz, V.M. Chaotically Accelerated Polymerase Chain Reaction by Microscale Rayleigh-Benard Convection. *Angew. Chem. Int. Edit.* **50**, 3048-3052 (2011).

121.   Agrawal, N., Hassan, Y.A. & Ugaz, V.M. A pocket-sized convective PCR thermocycler. *Angew. Chem. Int. Edit.* **46**, 4316-4319 (2007).

122.   Bailey, R.C., Kwong, G.A., Radu, C.G., Witte, O.N. & Heath, J.R. DNA-Encoded Antibody Libraries: A Unified Platform for Multiplexed Cell Sorting and Detection of Genes and Proteins. *J. Am. Chem. Soc.* **129**, 1959-1967 (2007).

123. Qavi, A.J. & Bailey, R.C. Multiplexed Detection and Label-Free Quantitation of MicroRNAs Using Arrays of Silicon Photonic Microring Resonators. *Angew. Chem. Int. Edit.* **49**, 4608-4611 (2010).

124. Hartman, M.R. et al. Thermostable Branched DNA Nanostructures as Modular Primers for Polymerase Chain Reaction. *Angew. Chem. Int. Edit.* **52**, 8699-8702 (2013).

125. Kim, J., Johnson, M., Hill, P. & Gale, B.K. Microfluidic sample preparation: cell lysis and nucleic acid purification. *Integr. Biol.* **1**, 574-586 (2009).

126. Chen, L., Manz, A. & Day, P.J.R. Total nucleic acid analysis integrated on microfluidic devices. *Lab Chip* **7**, 1413-1423 (2007).

127. Wilson, I.G. Inhibition and facilitation of nucleic acid amplification. *Appl. Environ. Microb.* **63**, 3741-3751 (1997).

128. Rådström, P., Knutsson, R., Wolffs, P., Lövenklev, M. & Löfström, C. Pre-PCR processing. *Mol. Biotechnol.* **26**, 133-146 (2004).

129. Erlich, H.A. PCR technology. Principles and applications for DNA amplification. (Stockton press, 1989).

130. Trombley Hall, A., McKay Zovanyi, A., Christensen, D.R., Koehler, J.W. & Devins Minogue, T. Evaluation of Inhibitor-Resistant Real-Time PCR Methods for Diagnostics in Clinical and Environmental Samples. *PLoS One* **8**, e73845 (2013).

131. Monteiro, L. et al. Complex polysaccharides as PCR inhibitors in feces: Helicobacter pylori model. *J. Clin. Microbiol.* **35**, 995-998 (1997).

132. Al-Soud, W.A. & Rådström, P. Purification and characterization of PCR-inhibitory components in blood cells. *J. Clin. Microbiol.* **39**, 485-493 (2001).

133. Legendre, L.A., Bienvenue, J.M., Roper, M.G., Ferrance, J.P. & Landers, J.P. A simple, valveless microfluidic sample preparation device for extraction and amplification of DNA from nanoliter-volume samples. *Anal. Chem.* **78**, 1444-1451 (2006).

134. Hong, J.W., Studer, V., Hang, G., Anderson, W.F. & Quake, S.R. A nanoliter-scale nucleic acid processor with parallel architecture. *Nat. Biotechnol.* **22**, 435-439 (2004).

135. Toriello, N.M. et al. Integrated microfluidic bioprocessor for single-cell gene expression analysis. *P. Natl. Acad. Sci. USA* **105**, 20173-20178 (2008).

136. White, A.K. et al. High-throughput microfluidic single-cell RT-qPCR. *P. Natl. Acad. Sci. USA* **108**, 13999-14004 (2011).

137. Lien, K.-Y., Lee, W.-C., Lei, H.-Y. & Lee, G.-B. Integrated reverse transcription polymerase chain reaction systems for virus detection. *Biosens. Bioelectron.* **22**, 1739-1748 (2007).

138. Gilbert, T.W., Sellaro, T.L. & Badylak, S.F. Decellularization of tissues and organs. *Biomaterials* **27**, 3675-3683 (2006).

139. Manage, D.P. et al. On-chip PCR amplification of genomic and viral templates in unprocessed whole blood. *Microfluidic. Nanofluidic.* **10**, 697-702 (2011).

140. Mao, F., Leung, W.-Y. & Xin, X. Characterization of EvaGreen and the implication of its physicochemical properties for qPCR applications. *BMC Biotechnol.* **7**, 76 (2007).

141. Wang, Y. et al. A novel strategy to engineer DNA polymerases for enhanced processivity and improved performance in vitro. *Nucleic Acids Res.* **32**, 1197-1207 (2004).

142. Novak, R. et al. Single-Cell Multiplex Gene Detection and Sequencing with Microfluidically Generated Agarose Emulsions. *Angew. Chem. Int. Edit.* **50**, 390-395 (2011).

143. Liu, P., Li, X., Greenspoon, S.A., Scherer, J.R. & Mathies, R.A. Integrated DNA purification, PCR, sample cleanup, and capillary electrophoresis microchip for forensic human identification. *Lab Chip* **11**, 1041-1048 (2011).

144. Leung, K. et al. A programmable droplet-based microfluidic device applied to multiparameter analysis of single microbes and microbial communities. *P. Natl. Acad. Sci. USA* **109**, 7665-7670 (2012).

145. Ma, S. et al. Diffusion-based microfluidic PCR for "one-pot" analysis of cells. *Lab Chip* **14**, 2905-2909 (2014).

146. Thorsen, T., Maerkl, S.J. & Quake, S.R. Microfluidic large-scale integration. *Science* **298**, 580-584 (2002).

147. Unger, M.A., Chou, H.-P., Thorsen, T., Scherer, A. & Quake, S.R. Monolithic Microfabricated Valves and Pumps by Multilayer Soft Lithography. *Science* **288**, 113-116 (2000).

148. Heyries, K.A. et al. Megapixel digital PCR. *Nat. Methods* **8**, 649-651 (2011).

149. Lukacs, G.L. et al. Size-dependent DNA mobility in cytoplasm and nucleus. *J. Biol. Chem.* **275**, 1625-1629 (2000).

150. Bu, Z., Biehl, R., Monkenbusch, M., Richter, D. & Callaway, D.J. Coupled protein domain motion in Taq polymerase revealed by neutron spin-echo spectroscopy. *P. Natl. Acad. Sci. USA* **102**, 17646-17651 (2005).

151. Nenninger, A., Mastroianni, G. & Mullineaux, C.W. Size dependence of protein diffusion in the cytoplasm of Escherichia coli. *J. Bacteriol.* **192**, 4535-4540 (2010).

152. Mamedov, T. et al. A fundamental study of the PCR amplification of GC-rich DNA templates. *Comput. Biol. Chem.* **32**, 452-457 (2008).

153. Hubley, M.J., Locke, B.R. & Moerland, T.S. The effects of temperature, pH, and magnesium on the diffusion coefficient of ATP in solutions of physiological ionic strength. *Biochim. Biophys. Acta* **1291**, 115-121 (1996).

154. Leaist, D.G. A moving-boundary technique for the measurement of diffusion in liquids. triton X-100 in water. *J. Solution Chem.* **20**, 187-197 (1991).

155. Yuan-Hui, L. & Gregory, S. Diffusion of ions in sea water and in deep-sea sediments. *Geochim. Cosmochim. Ac.* **38**, 703-714 (1974).

156. Hansen, K.D. et al. Increased methylation variation in epigenetic domains across cancer types. *Nat. Genet.* **43**, 768-775 (2011).

157. Spiers, H. et al. Methylomic trajectories across human fetal brain development. *Genome Res.* **25**, 338-352 (2015).

158. Numata, S. et al. DNA methylation signatures in development and aging of the human prefrontal cortex. *Am. J. Hum. Genet.* **90**, 260-272 (2012).

159. Mo, A. et al. Epigenomic Signatures of Neuronal Diversity in the Mammalian Brain. *Neuron* **86**, 1369-1384 (2015).

160. Cokus, S.J. et al. Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature* **452**, 215-219 (2008).

161. Serre, D., Lee, B.H. & Ting, A.H. MBD-isolated Genome Sequencing provides a high-throughput and comprehensive survey of DNA methylation in the human genome. *Nucleic Acids Res.* **38**, 391-399 (2010).

162. Gu, H. et al. Genome-scale DNA methylation mapping of clinical samples at single-nucleotide resolution. *Nat. Methods* **7**, 133-136 (2010).

163. Adey, A. & Shendure, J. Ultra-low-input, tagmentation-based whole-genome bisulfite sequencing. *Genome Res.* **22**, 1139-1143 (2012).

164. Farlik, M. et al. Single-Cell DNA Methylome Sequencing and Bioinformatic Inference of Epigenomic Cell-State Dynamics. *Cell Rep.* **10**, 1386-1397 (2015).

165. Ehrich, M., Zoll, S., Sur, S. & van den Boom, D. A new method for accurate assessment of DNA quality after bisulfite treatment. *Nucleic Acids Res.* **35**, e29 (2007).

166. Grunau, C., Clark, S.J. & Rosenthal, A. Bisulfite genomic sequencing: systematic investigation of critical experimental parameters. *Nucleic Acids Res.* **29**, e65 (2001).

167. Munson, K., Clark, J., Lamparska-Kupsik, K. & Smith, S.S. Recovery of bisulfite-converted genomic sequences in the methylation-sensitive QPCR. *Nucleic Acids Res.* **35**, 2893-2903 (2007).

168. Boyle, P. et al. Gel-free multiplexed reduced representation bisulfite sequencing for large-scale DNA methylation profiling. *Genome Biol.* **13**, R92 (2012).

169.   Smith, Z.D. et al. DNA methylation dynamics of the human preimplantation embryo. *Nature* **511**, 611-615 (2014).

170.   Reizel, Y. et al. Gender-specific postnatal demethylation and establishment of epigenetic memory. *Genes Dev.* **29**, 923-933 (2015).

171.   Gabel, H.W. et al. Disruption of DNA methylation-dependent long gene repression in Rett syndrome. *Nature* **522**, 89-93 (2015).

172.   Kaeser, P.S., Deng, L., Fan, M. & Südhof, T.C. RIM genes differentially contribute to organizing presynaptic release sites. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 11830-11835 (2012).

173.   Mikkelsen, T.S. et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553-560 (2007).

174.   Davies, M.N. et al. Functional annotation of the human brain methylome identifies tissue-specific epigenetic variation across brain and blood. *Genome Biol.* **13**, R43 (2012).

175.   Meissner, A. et al. Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* **454**, 766-770 (2008).

176.   Angermueller, C. et al. Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. *Nat. Methods* **13**, 229-232 (2016).

177.   Unger, M.A., Chou, H.P., Thorsen, T., Scherer, A. & Quake, S.R. Monolithic microfabricated valves and pumps by multilayer soft lithography. *Science* **288**, 113-116 (2000).

178.   Lake, B.B. et al. Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. *Science* **352**, 1586-1590 (2016).

179. Lister, R. et al. Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature* **471**, 68-73 (2011).

180. Krueger, F. & Andrews, S.R. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27**, 1571-1572 (2011).

181. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S.L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).

182. Akalin, A. et al. methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol.* **13**, 2012-2013 (2012).

183. Robinson, J.T. et al. Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24-26 (2011).

184. Wu, H., Wang, C. & Wu, Z. A new shrinkage estimator for dispersion improves differential expression detection in RNA-seq data. *Biostatistics* **14**, 232-243 (2013).

185. Dennis, G. et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol.* **4**, R60 (2003).

186. Trapnell, C., Pachter, L. & Salzberg, S.L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105-1111 (2009).

187. Trapnell, C. et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511-515 (2010).

188. Samson, E., Marchand, J. & Snyder, K.A. Calculation of ionic diffusion coefficients on the basis of migration test results. *Mater. Struct.* **36**, 156-165 (2003).

189. Hayduk, W. & Laudie, H. Prediction of diffusion coefficients for nonelectrolytes in dilute aqueous solutions. *AIChE J.* **20**, 611-615 (1974).

190. Nkodo, A.E. et al. Diffusion coefficient of DNA molecules during free solution electrophoresis. *Electrophoresis* **22**, 2424-2432 (2001).

191. Fan, R. et al. Integrated barcode chips for rapid, multiplexed analysis of proteins in microliter quantities of blood. *Nat. Biotechnol.* **26**, 1373-1378 (2008).

192. Shankaranarayanan, P. et al. Single-tube linear DNA amplification (LinDA) for robust ChIP-seq. *Nat. Methods* **8**, 565-U565 (2011).

193. Yamamoto, Y., Verma, U.N., Prajapati, S., Kwak, Y.T. & Gaynor, R.B. Histone H3 phosphorylation by IKK-alpha is critical for cytokine-induced gene expression. *Nature* **423**, 655-659 (2003).

194. Li, J. et al. EZH2-mediated H3K27 trimethylation mediates neurodegeneration in ataxia-telangiectasia. *Nat. Neurosci.* **16**, 1745-1753 (2013).

195. Cheung, I. et al. Developmental regulation and individual differences of neuronal H3K4me3 epigenomes in the prefrontal cortex. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 8824-8829 (2010).

196. Voigt, P., Tee, W.-W. & Reinberg, D. A double take on bivalent promoters. *Genes Dev.* **27**, 1318-1338 (2013).

197. Philippidou, P. & Dasen, J.S. Hox genes: choreographers in neural development, architects of circuit organization. *Neuron* **80**, 12-34 (2013).

198. Frank, C.L. et al. Regulation of chromatin accessibility and Zic binding at enhancers in the developing cerebellum. *Nat. Neurosci.* **18**, 647-656 (2015).

199. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).

200. Xu, S., Grullon, S., Ge, K. & Peng, W. Spatial Clustering for Identification of ChIP-Enriched Regions (SICER) to Map Regions of Histone Methylation Patterns in Embryonic Stem Cells. *Methods Mol. Biol.* **1150**, 97-111 (2014).

201. Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

202. Firpi, H.A., Ucar, D. & Tan, K. Discover regulatory DNA elements using chromatin signatures and artificial neural network. *Bioinformatics* **26**, 1579-1586 (2010).

203. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576-589 (2010).

204. Zhuang, G., Katakura, Y., Omasa, T., Kishimoto, M. & Suga, K. Measurement of association rate constant of antibody-antigen interaction in solution based on enzyme-linked immunosorbent assay. *J. Biosci. Bioeng.* **92**, 330-336 (2001).

205. Jakobsen, J.S. et al. Amplification of pico-scale DNA mediated by bacterial carrier DNA for small-cell-number transcription factor ChIP-seq. *BMC Genomics* **16**, 46 (2015).

206. Wang, Z., Gerstein, M. & Snyder, M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* **10**, 57-63 (2009).

207. Schmittgen, T.D. & Livak, K.J. Analyzing real-time PCR data by the comparative CT method. *Nat. Protoc.* **3**, 1101-1108 (2008).

208. Sealfon, S.C. & Chu, T.T. RNA and DNA microarrays. *Methods Mol. Biol.* **671**, 3-34 (2011).

209. Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* **5**, 621-628 (2008).

210. Croucher, N.J. & Thomson, N.R. Studying bacterial transcriptomes using RNA-seq. *Curr. Opin. Microbiol.* **13**, 619-624 (2010).

211. Fleige, S. & Pfaffl, M.W. RNA integrity and the effect on the real-time qRT-PCR performance. *Mol. Aspects Med.* **27**, 126-139 (2006).

212. Imbeaud, S. et al. Towards standardization of RNA quality assessment using user-independent classifiers of microcapillary electrophoresis traces. *Nucleic Acids Res.* **33** (2005).

213. Hurley, S.S., Splitter, G. & Welch, R. Rapid lysis technique for mycobacterial species. *J. Clin. Microbiol.* **25**, 2227-2229 (1987).

214. Rio, D.C., Ares, M., Jr., Hannon, G.J. & Nilsen, T.W. Purification of RNA using TRIzol (TRI reagent). *Cold Spring Harb. Protoc.* **6** (2010).

215. Mangan, J.A., Sole, K.M., Mitchison, D.A. & Butcher, P.D. An effective method of RNA extraction from bacteria refractory to disruption, including mycobacteria. *Nucleic Acids Res.* **25**, 675-676 (1997).

216. Aldous, W.K., Pounder, J.I., Cloud, J.L. & Woods, G.L. Comparison of Six Methods of Extracting Mycobacterium tuberculosis DNA from Processed Sputum for Testing by Quantitative Real-Time PCR. *J. Clin. Microbiol.* **43**, 2471-2473 (2005).

217. Jahn, C.E., Charkowski, A.O. & Willis, D.K. Evaluation of isolation methods and RNA integrity for bacterial RNA quantitation. *J. Microbiol. Meth.* **75**, 318-324 (2008).

218. Vandeventer, P.E. et al. Mechanical Disruption of Lysis-Resistant Bacterial Cells by Use of a Miniature, Low-Power, Disposable Device. *J. Clin. Microbiol.* **49**, 2533-2539 (2011).

219. Leite, G.M., Magan, N. & Medina, A. Comparison of different bead-beating RNA extraction strategies: an optimized method for filamentous fungi. *J. Microbiol. Methods* **88**, 413-418 (2012).

220. Stahl, D.A. & Urbance, J.W. The division between fast- and slow-growing species corresponds to natural relationships among the mycobacteria. *J. Bacteriol.* **172**, 116-124 (1990).

221. Wang, H.-Y., Bhunia, A.K. & Lu, C. A microfluidic flow-through device for high throughput electrical lysis of bacterial cells based on continuous dc voltage. *Biosens. Bioelectron.* **22**, 582-588 (2006).

222. Bao, N. & Lu, C. A microfluidic device for physical trapping and electrical lysis of bacterial cells. *Appl. Phys. Lett.* **92** (2008).

223. Vulto, P. et al. A microfluidic approach for high efficiency extraction of low molecular weight RNA. *Lab Chip* **10**, 610-616 (2010).

224. Ma, S. et al. Electroporation-based delivery of cell-penetrating peptide conjugates of peptide nucleic acids for antisense inhibition of intracellular bacteria. *Integr. Biol.* **6**, 973-978 (2014).

225. Sun, C., Cao, Z., Wu, M. & Lu, C. Intracellular Tracking of Single Native Molecules with Electroporation-Delivered Quantum Dots. *Anal. Chem.* **86**, 11403-11409 (2014).

226. Bao, N., Wang, J. & Lu, C. Microfluidic electroporation for selective release of intracellular molecules at the single-cell level. *Electrophoresis* **29**, 2939-2944 (2008).

227. Wong, M.L. & Medrano, J.F. Real-time PCR for mRNA quantitation. *Biotechniques* **39**, 75-85 (2005).

228. Lundberg, P. & Langel, U. A brief introduction to cell-penetrating peptides. *J. Mol. Recognit.* **16**, 227-233 (2003).

229. Madani, F., Lindberg, S., Langel, U., Futaki, S. & Graslund, A. Mechanisms of cellular uptake of cell-penetrating peptides. *J. Biophys.* **2011**, 414729 (2011).

230. Berry, C.C. Intracellular delivery of nanopartides via the HIV-1 tat pepticle. *Nanomedicine* **3**, 357-365 (2008).

231. Vives, E., Brodin, P. & Lebleu, B. A truncated HIV-1 Tat protein basic domain rapidly translocates through the plasma membrane and accumulates in the cell nucleus. *J. Biol. Chem.* **272**, 16010-16017 (1997).

232. Silhol, M., Tyagi, M., Giacca, M., Lebleu, B. & Vives, E. Different mechanisms for cellular internalization of the HIV-1 Tat-derived cell penetrating peptide and recombinant proteins fused to Tat. *Eur. J. Biochem.* **269**, 494-501 (2002).

233. Pooga, M., Hallbrink, M., Zorko, M. & Langel, U. Cell penetration by transportan. *FASEB J.* **12**, 67-77 (1998).

234. Fisher, L. et al. Cellular delivery of a double-stranded oligonucleotide NF kappa B decoy by hybridization to complementary PNA linked to a cell-penetrating peptide. *Gene Ther.* **11**, 1264-1272 (2004).

235. Nekhotiaeva, N. et al. Cell entry and antimicrobial properties of eukaryotic cell-penetrating peptides. *FASEB J.* **18**, 394-396 (2004).

236. Anko, M. et al. Influence of stearyl and trifluoromethylquinoline modifications of the cell penetrating peptide TP10 on its interaction with a lipid membrane. *BBA-Biomembranes* **1818**, 915-924 (2012).

237. Arrighi, R.B.G. et al. Cell-penetrating peptide TP10 shows broad-spectrum activity against both Plasmodium falciparum and Trypanosoma brucei brucei. *Antimicrob. Agents Ch.* **52**, 3414-3417 (2008).

238. Pooga, M. et al. Cell penetrating PNA constructs regulate galanin receptor levels and modify pain transmission in vivo. *Nat. Biotechnol.* **16**, 857-861 (1998).

239. Elmquist, A., Lindgren, M., Bartfai, T. & Langel, U. VE-cadherin-derived cell-penetrating peptide, pVEC, with carrier functions. *Exp. Cell Res.* **269**, 237-244 (2001).

240. Elmquist, A. & Langel, U. In vitro uptake and stability study of pVEC and its all-D analog. *Biol. Chem.* **384**, 387-393 (2003).

241. Egholm, M. et al. PNA hybridizes to complementary oligonucleotides obeying the Watson-Crick hydrogen-bonding rules. *Nature* **365**, 566-568 (1993).

242. Shakeel, S., Karim, S. & Ali, A. Peptide nucleic acid (PNA)—a review. *J. Chem. Technol. Biot.* **81**, 892-899 (2006).

243. Crooke, S.T. Antisense drug technology: principles, strategies, and applications. (CRC Press, 2010).

244. Turner, J.J. et al. Cell-penetrating peptide conjugates of peptide nucleic acids (PNA) as inhibitors of HIV-1 Tat-dependent trans-activation in cells. *Nucleic Acids Res.* **33**, 6837-6849 (2005).

245. Mae, M., El Andaloussi, S., Lehto, T. & Langel, U. Chemically modified cell-penetrating peptides for the delivery of nucleic acids. *Expert Opin. Drug Deliv.* **6**, 1195-1205 (2009).

246. Lundberg, P., El-Andaloussi, S., Sutlu, T., Johansson, H. & Langel, U. Delivery of short interfering RNA using endosomolytic cell-penetrating peptides. *FASEB J.* **21**, 2664-2671 (2007).

247. Salomone, F. et al. A novel chimeric cell-penetrating peptide with membrane-disruptive properties for efficient endosomal escape. *J. Control. Release* **163**, 293-303 (2012).

248. Shiraishi, T. & Nielsen, P.E. Enhanced delivery of cell-penetrating peptide-peptide nucleic acid conjugates by endosomal disruption. *Nat. Protoc.* **1**, 633-636 (2006).

249. Hatamoto, M., Ohashi, A. & Imachi, H. Peptide nucleic acids (PNAs) antisense effect to bacterial growth and their application potentiality in biotechnology. *Appl. Microbiol. Biot.* **86**, 397-402 (2010).

250. Bai, H. et al. Targeting RNA Polymerase Primary σ 70 as a Therapeutic Strategy against Methicillin-Resistant Staphylococcus aureus by Antisense Peptide Nucleic Acid. *PLoS One* **7**, e29886 (2012).

251. Bai, H. et al. Antisense antibiotics: a brief review of novel target discovery and delivery. *Curr. Drug Discovery Technol.* **7**, 76-85 (2010).

252. Good, L., Awasthi, S.K., Dryselius, R., Larsson, O. & Nielsen, P.E. Bactericidal antisense effects of peptide–PNA conjugates. *Nat. Biotechnol.* **19**, 360-364 (2001).

253. Good, L., Sandberg, R., Larsson, O., Nielsen, P.E. & Wahlestedt, C. Antisense PNA effects in Escherichia coli are limited by the outer-membrane LPS layer. *Microbiology* **146**, 2665-2670 (2000).

254. Good, L. & Nielsen, P.E. Inhibition of translation and bacterial growth by peptide nucleic acid targeted to ribosomal RNA. *P. Natl. Acad. Sci. USA* **95**, 2073-2076 (1998).

255. Vaara, M. & Porro, M. Group of peptides that act synergistically with hydrophobic antibiotics against gram-negative enteric bacteria. *Antimicrob. Agents Ch.* **40**, 1801-1805 (1996).

256. Bendifallah, N. et al. Evaluation of cell-penetrating peptides (CPPs) as vehicles for intracellular delivery of antisense peptide nucleic acid (PNA). *Bioconjugate Chem.* **17**, 750-758 (2006).

257. Galan, J.E. & Zhou, D. Striking a balance: Modulation of the actin cytoskeleton by Salmonella. *P. Natl. Acad. Sci. USA* **97**, 8754-8761 (2000).

258. Soofi, M.A. & Seleem, M.N. Targeting Essential Genes in Salmonella enterica Serovar Typhimurium with Antisense Peptide Nucleic Acid. *Antimicrob. Agents Ch.* **56**, 6407-6409 (2012).

259. Abes, S. et al. Endosome trapping limits the efficiency of splicing correction by PNA-oligolysine conjugates. *J. Control. Release* **110**, 595-604 (2006).

260. Nielsen, P.E. Addressing the challenges of cellular delivery and bioavailability of peptide nucleic acids (PNA). *Q. Rev. Biophys.* **38**, 345-350 (2005).

261. Tsolis, R.M., Bäumler, A.J. & Heffron, F. Role of Salmonella typhimurium Mn-superoxide dismutase (SodA) in protection against early killing by J774 macrophages. *Infect. Immun.* **63**, 1739-1744 (1995).

262. Lindgren, S.W., Stojiljkovic, I. & Heffron, F. Macrophage killing is an essential virulence mechanism of Salmonella typhimurium. *P. Natl. Acad. Sci. USA* **93**, 4197-4201 (1996).

263. Buchmeier, N.A. & Heffron, F. Inhibition of macrophage phagosome-lysosome fusion by Salmonella typhimurium. *Infect. Immun.* **59**, 2232-2238 (1991).

264. Rengarajan, J., Bloom, B.R. & Rubin, E.J. Genome-wide requirements for Mycobacterium tuberculosis adaptation and survival in macrophages. *P. Natl. Acad. Sci. USA* **102**, 8327-8332 (2005).

265. Ernst, R.K., Guina, T. & Miller, S.I. How Intracellular Bacteria Survive: Surface Modifications That Promote Resistance to Host Innate Immune Responses. *J. Infect. Dis.* **179**, S326-S330 (1999).

266. Nau, G.J. et al. Human macrophage activation programs induced by bacterial pathogens. *P. Natl. Acad. Sci. USA* **99**, 1503-1508 (2002).

267. Graham, J.E. & Clark-Curtiss, J.E. Identification of Mycobacterium tuberculosis RNAs synthesized in response to phagocytosis by human macrophages by selective capture of transcribed sequences (SCOTS). *P. Natl. Acad. Sci. USA* **96**, 11554-11559 (1999).

268. Rols, M.P. et al. In vivo electrically mediated protein and gene transfer in murine melanoma. *Nat. Biotechnol.* **16**, 168-171 (1998).

269. Wang, J., Zhan, Y., Ugaz, V.M. & Lu, C. Vortex-assisted DNA delivery. *Lab Chip* **10**, 2057-2061 (2010).

270. Geng, T., Zhan, Y.H., Wang, J. & Lu, C. Transfection of cells using flow-through electroporation based on constant voltage. *Nat. Protoc.* **6**, 1192-1208 (2011).

271. Seleem, M.N. et al. Targeting Brucella melitensis with polymeric nanoparticles containing streptomycin and doxycycline. *FEMS Microbiol. Lett.* **294**, 24-31 (2009).

272. Zhan, Y. et al. Release of Intracellular Proteins by Electroporation with Preserved Cell Viability. *Anal. Chem.* **84**, 8102-8105 (2012).

273. Wang, F. et al. Microfluidic delivery of small molecules into mammalian cells based on hydrodynamic focusing. *Biotechnol. Bioeng.* **100**, 150-158 (2008).

274. Teissie, J., Golzio, M. & Rols, M.P. Mechanisms of cell membrane electropermeabilization: a minireview of our present knowledge. *Biochim. Biophys. Acta* **1724**, 270-280 (2005).

275. Weaver, J.C. & Chizmadzhev, Y.A. Theory of electroporation: A review. *Bioelectrochem. Bioenerg.* **41**, 135-160 (1996).

276. Yu, M., Tan, W.C. & Lin, H. A stochastic model for DNA translocation through an electropore. *BBA-Biomembranes* **1818**, 2494-2501 (2012).

277. Li, J.B. & Lin, H. Numerical simulation of molecular uptake via electroporation. *Bioelectrochemistry* **82**, 10-21 (2011).

278. Wang, H.Y., Bhunia, A.K. & Lu, C. A microfluidic flow-through device for high throughput electrical lysis of bacterial cells based on continuous dc voltage. *Biosens. Bioelectron.* **22**, 582-588 (2006).

PUBLICATIONS

Journal Papers

1. **Ma, S.** & Lu, C. Paralleled Microfluidic Device for High-throughput and Low-input Histone Modification Analysis Based on Surface ChIP (SurfaceChIP-seq). In preparation

2. **Ma, S.**, Sun, Z., Sun, C., Murphy, T. W., Xie, H. & Lu. C. Cell Type-specific Mouse Brain Methylomes Profiled Using an Ultralow-input Microfluidic Device. ***Nat. Biomed. Eng.*** under review.

3. Murphy, T. W., **Ma, S.** & Lu, C. Microfluidics for single cell analysis. In preparation

4. Murphy, T. W., **Ma, S.**, Cao, Z. & Lu, C. Fluidized chromatin immunoprecipitation for automated histone modification profiling. In preparation.

5. Sun, C., **Ma, S.** & Lu, C. On Chip Chromosome Conformation Capture (3C) Technology for Examining Genome-Wide Chromatin Interactions with High Sensitivity. In preparation.

6. **Ma, S.**, Murphy, T. W. & Lu, C. Microfluidics for Genome-wide Studies Involving Next Generation Sequencing. (review) 11, 021501 (2017) ***Biomicrofluidics*** (impact factor: 2.708)

7. **Ma, S.**, Bryson, B.D., Sun, C., Fortune, S.M. & Lu, C. RNA extraction from a mycobacterium under ultrahigh electric field intensity. ***Anal. Chem.*** 88, 5053-5057 (2016). (impact factor: 5.886)

8. **Ma, S.**, Loufakis, D. N., Cao Z., Chang Y., Achenie, L. EK & Lu C. Diffusion-based microfluidic PCR for "one-pot" analysis of cells. *Lab Chip* 14, 2905-2909 (2014). (impact factor: 5.586)

9. **Ma, S.**, Schroeder, B., Sun, C., Loufakis, D. N., Cao, Z. Sriranganathan, N. & Lu, C. Electroporation-based delivery of cell-penetrating peptide conjugates of peptide nucleic acids for antisense inhibition of intracellular bacteria. *Integr. Biol.* 2014, 973-978 (2014). (impact factor: 3.371)

10. **Ma, S.**, Tang, Y., Liu, J. & Wu, J. Visible paper chip immunoassay for rapid determination of bacteria in water distribution system. *Talanta* 120, 135-140 (2014). (impact factor: 4.035)

11. Sun, C., Hsieh, Y., **Ma, S.**, Geng, S., Cao, Z., Li, L. & Lu, C. Immunomagnetic separation of tumor initiating cells by screening two surface markers. *Sci. Rep.* 7 (2017). (impact factor: 5.228)

12. Sun, C., Hassanisaber, H., Yu, R., **Ma, S.**, Verbridge, S. S. & Lu, C. Paramagnetic structures within a microfluidic channel for enhanced immunomagnetic isolation and surface patterning of cells. *Sci. Rep.* 6 (2016). (impact factor: 5.228)

13. Loufakis, D.N., Cao, Z., **Ma, S.,** Mittelman, D. & Lu, C. Focusing of mammalian cells under an ultrahigh pH gradient created by unidirectional electropulsation in a confined microchamber. *Chem. Sci.* 5, 3331-3337 (2014). (impact factor: 9.144)

14. Sun, C., Ouyang M., Cao, Z., **Ma, S.**, Alqublan, H., Sriranganathan, N., Wang, Y. & Lu, C. Electroporation-delivered fluorescent protein biosensors for probing molecular activities in cells without genetic encoding. *Chem. Comm.* 50, 11536-11539 (2014). (impact factor: 6.567)

15. Tan, J., Zhao W., Yu J., **Ma, S.**, Sailor, M. J. & Wu, J. Capture, enrichment, and mass spectrometric detection of low-molecular-weight biomarkers with nanoporous silicon microparticles. ***Adv. Healthc. Mater.*** 1, 742-750 (2012). (impact factor: 5.760)

16. Liu, Z., Wu, Q., Song, X. & **Ma, S.** Solid high-proton conductors based on heteropoly acids. ***Prog. Chem.*** 21, 982-989 (2009). (impact factor: 0.894)

Conferences

1. **Ma, S.** & Lu, C Microfluidics for low-input mapping of epigenomes. 2017 AACR Annual Meeting, Apr 1-5 2017, Washington, D.C. (USA) (Oral presentation)

2. **Ma, S.**, Sun, Z., Sun, C., Murphy, T. W., Xie, H. & Lu, C. Ultrasensitive microfluidic assay for genome-wide DNA methylation analysis and precision medicine. 2016 AIChE Annual Meeting, Nov 13-18 2016, San Francisco, CA (USA) (Oral presentation)

3. **Ma, S.**, Luo, C., Sun, Z., Lucero, J., Sun, C., Murphy, T. W., Xie, H., Ecker J. R., Behrens, M. M. & Lu, C. A Microfluidic Device for Low-Input Methylomic Analysis Based on Reduced Representative Bisulfite Sequencing. 2016 AIChE Annual Meeting, Nov 13-18 2016, San Francisco, CA (USA) (Poster presentation)

4. Sun, C., Hassanisaber, H., Yu, R., **Ma, S.**, Verbridge, S. & Lu, C. Enhanced microfluidic immunomagnetic separation based on microfabricated patterns from paramagnetic nanoparticles. 2015 AIChE Annual Meeting, Nov 8-13 2015, Salt Lake City, UT (USA)

5. **Ma, S.**, Luo, C., Sun, Z., Lucero, J., Sun, C., Murphy, T. W., Xie, H., Ecker J. R., Behrens, M. M. & Lu, C. Ultrasensitive microfluidic assay for genome-wide DNA methylation analysis and precision medicine. 2016 BMES Annual Meeting, Oct 11-14 2016, San Francisco, CA (USA) (Poster presentation)

6. **Ma, S.**, Bryson, B. D., Sun, C., Fortune, S. & Lu, C. RNA extraction from a mycobacterium under ultrahigh electric field intensity in a microfluidic device. 2016 BMES Annual Meeting, Oct 11-14 2016, San Francisco, CA (USA) (Oral presentation)

7. **Ma, S.**, Sun, Z., Murphy, T. W., Xie, H., Ecker J. R., Behrens, M. M. & Lu, C. Ultrasensitive Microfluidic Assays for Epigenomic Analysis of Neurons from

Mammalian Brains. 15th SBES Annual Symposium, May 11 2016, Winston-Salem, NC (USA) (Oral presentation)

8. **Ma, S**. & Lu, C. Diffusion-based microfluidic bisulfite conversion for DNA methylation detection. 2015 AIChE Annual Meeting, Nov 8-13 2015, Salt Lake City, UT (USA) (Oral presentation)

9. **Ma, S.**, Bryson, B. D., Sun, C., Fortune, S. & Lu, C. mRNA extraction from mycobacteria M. Smegmatis utilizing ultrahigh field intensity electrolysis. 2015 AIChE Annual Meeting, Nov 8-13 2015, Salt Lake City, UT (USA) (Oral presentation)

10. Sun, C., Hassanisaber, H., Yu, R., **Ma, S.**, Verbridge, S. & Lu, C. Enhanced Microfluidic Immunomagnetic Separation Based on Microfabricated paramagnetic Patterns. 2015 BMES Annual Meeting, Oct 7-10 2015, Tampa, FL (USA)

11. **Ma, S.** & Lu, C. Diffusion-based ultrasensitive bisulfite conversion on chip. 14th SBES Annual Symposium, May 13 2015, Blacksburg, VA (USA) (Poster presentation)

12. **Ma, S.**, Loufakis, D. N., Cao Z., Chang Y., Achenie, L. EK & Lu C. Diffusion-based Microfluidic PCR for "One-pot" Analysis of Cells. 2014 AIChE Annual Meeting, Nov 16-21 2014, Atlanta, GA (USA) (Oral presentation)

13. **Ma, S.**, Schroeder, B., Sun, C., Loufakis, D. N., Cao, Z. Sriranganathan, N., & Lu, C. Microfluidic electroporation for delivery of cell-penetrating peptide conjugates of Peptide Nucleic Acids (PNA) for antisense inhibition of intracellular bacteria. 2014 AIChE Annual Meeting, Nov 16-21 2014, Atlanta, GA (USA) (Oral presentation)

14. Sun, C., Wang, Y., Ouyang, M., **Ma, S.**, Wang, Y. & Lu, C. Electroporation-delivered protein biosensors for study of molecular activity. 2014 AIChE Annual Meeting, Nov 16-21 2014, Atlanta, GA (USA)

15. **Ma, S.**, Loufakis, D. N., Cao Z., Chang Y., Achenie, L. EK & Lu C. Diffusion-based Microfluidic PCR for "One-pot" Analysis of Cells. 18th International Conference on Miniaturized Systems for Chemistry and Life Sciences (µTAS), Oct 26-30 2014, San Antonio, TX (USA) (Poster presentation)

16. **Ma, S.**, Schroeder, B., Sun, C., Loufakis, D. N., Cao, Z. Sriranganathan, N. & Lu, C. Microfluidic electroporation for delivery of cell-penetrating peptide conjugates of Peptide Nucleic Acids (PNA) for antisense inhibition of intracellular bacteria. 18th International Conference on Miniaturized Systems for Chemistry and Life Sciences (µTAS), Oct 26-30 2014, San Antonio, TX (USA) (Poster presentation)

17. Sun, C., Ouyang, M., Cao, Z., **Ma, S.**, Wang, Y. & Lu, C. Electroporation-delivered protein biosensors for study of molecular activity on microfluidic platform. 18th International Conference on Miniaturized Systems for Chemistry and Life Sciences (µTAS), Oct 26-30 2014, San Antonio, TX (USA)

18. Lu, C., **Ma, S.**, Loufakis, D. N., Cao Z., Chang Y. & Achenie, L. EK. Diffusion-based Microfluidic PCR for "One-pot" Analysis of Cells. 2014 BMES Annual Meeting, Oct 22-25 2016, San Antonio, TX (USA)

19. Loufakis, D. N., Cao, Z., **Ma, S.**, Mittelman, D. & Lu, C. Cell allignement under unidirectional electropulsation in a microfluidic device. AICHE 2013 3-8 Nov 2013, San Francisco, CA (USA)

20. Loufakis, D. N., Lu, C., Cao, Z., **Ma, S.** & Mittelman, D. Alignment of cells under unidirectional electric pulses., FACSS SCIX 2013 29 Sept- 4 Oct 2013, Milwaukee, WI (USA)

21. **Ma, S.**, Loufakis, D. N., Cao Z., Chang Y., Achenie, L. EK & Lu C. Diffusion-based Microfluidic PCR for "One-pot" Analysis of Cells. 13th SBES Annual Symposium, May 13 2014, Winston-Salem, NC (USA) (Poster presentation)

22. **Ma, S.**, Loufakis, D. N., Cao Z., Chang Y., Achenie, L. EK & Lu C. Electroporation enhances antisense Effect of peptide nucleic acids against intracellular bacteria. 12th SBES Annual Symposium, May 16 2014, Blacksburg, VA (USA) (Poster presentation)