Ph.D. DISSERTATION

High-throughput single molecule identification and separation method using focused pulse laser system

펄스 레이저를 이용한 고속 단분자 식별 및 분리 기법

BY

Howon Lee

AUGUST 2013

INTERDISCIPLINARY PROGRAM IN
NANO-SCIENCE AND TECHNOLOGY

COLLEGE OF NATURE SCIENCE
SEOUL NATIONAL UNIVERSITY

High-throughput single molecule identification and separation method using focused pulse laser system

펄스 레이저를 이용한 고속 단분자 식별 및 분리 기법

지도교수 권 성 훈

이 논문을 공학박사 학위논문으로 제출함

2013 년 6 월

서울대학교 대학원

협동과정 나노과학기술전공

이 호 원

이호원의 공학박사 학위논문을 인준함

2013 년 6 월

위 원 장 : 박 영 준

부위원장 : 권 성 훈

위 원 : 이 병 호

위 원 : 김 성 훈

위 원 : 방 두 희

Abstract

The key challenge of synthetic biology currently lies in the absence of cost effective high standard oligonucleotide precursor for constructing target long sequence. Microarray DNA is an ultra-rich source of oligonucleotides that generates millions of short oligonucleotide sequence in a single run. In spite of ensuring overwhelming advantages over conventional chemical oligonucleotide synthesis, the efficiency of the progress is dogged by high complexity and low quality of microarray DNA. In this thesis, I present various techniques including encoded-microparticle, DNA microarray and pulse laser sniper cloning, for the improvement of preparative tool for 'writing' DNA.

In the first part of this thesis, an important state-of-the-art element technologies for writing DNA are reviewed. Microarray DNA technology offers millions of short DNA in a cost effective single run that overcomes the problems related with conventional one-by-one column synthetic approach. Meanwhile, the downstream separation and the identification steps which normally consist a vector cloning and Sanger sequencing can also be replaced by high-throughput Next Generation Sequencing (NGS) platform. This chapter concludes in discussing the developmental possibility of a next generation writing technology by closely combining the elemental technologies into a preparative tool.

The second part provides an overview to the fabrication method of various

microparticles for complex pool separation. The microparticles which possess distinctive IDs and probe oligonucleotides on their surfaces plays a floating microfilter that selectively separates the target single strand DNA from complex pool according to the probe sequence. The fabrication of color barcoded microparticle and magnetochromatic sphere based on optofluidic technique is described followed by simple demonstration of DNA separation.

Third part describes the pulse laser driven microstructure techniques. The focused nanosecond pulse laser exerts radiation pressure onto the microparticles containing hybridization selected DNA from mixed pool. Furthermore, the target microparticles can be physically separated without actual physical contact. More condensed energy of focused pulse laser ablates target substrate and therefore, generates a small explosion. The separated contents of an array of probe spots such as DNA microarray is also able to be individualized for utilization by directly ablation of target containing substrate.

Chapter 4 presents the development of 'clone sniper' method using parallel identification followed by high-throughput separation approach to construct ultra-high quality oligonucleotide library with low cost and high-throughput. This approach reduces the labor intensive conventional clonal separation and expensive Sanger derived identification. The custom made pulse laser retrieval system enables non-contact contamination high-throughput separation of perfect parts from

sequencing plate with precise position data constructed by local mapping algorithm. The serial process consists of parallel synthesis parallel identification and high-throughput separation which not only increases the quality of contents, but also dramatically reduces the necessary resources, such as; cost, labor and time.

Chapter 5 provides a very compact summary of this research work, highlights the contributions made. Possibilities for future work to increase the significance of the approaches discussed.

Keywords: Microarray DNA, Encoded microparticle, Pulse laser, Next generation sequencing, Cloning.

Student Number: 2008-3013

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In this dissertation, I introduce a new platform for providing ultra-high quality oligonucleotide via combination of emerging idenfication technology and home-made pulse laser separation system. Parallel synthesis and parallel idenfication followed by high-throughput separation system dramatically lower down the cost, labor and maximize the utility of microarray DNA complex pool. The system presented here may have a significant impact on most of future biotechnology including functional genome study, synthetic biology and protein engineering with superior quality and throughput. In this chapter, several important related state-of-the-art element technologies for writing DNA are reviewed and examine the possibility of next generation writing technology by close combining those element technologies into preparative tool.

1.1 Synthetic DNA

For the last ten years, synthetic DNA attracts researchers' attention due to their accessibility and flexibility. Column based solid-phase synthesis has been widely used for constructing artificial DNA. Molecules are bound on a bead and synthesized step-by-step with sequential supply of reaction solution. Although in most service company supply artificial short length DNA with this method, relatively high cost, low throughput (~100 strand/run) and error prone process hinder the progress.

DNA chip

Eversince the introduction of first DNA chip (microarray) at Stanford university (Pat. Brown) in 1995, product has been commercialized and used for gene expression analysis, single nucleotide polymorphisms (SNPs) screening, pharmacogenomics search, infectious and genetic disease, cancer diagnostics and forensic and genetic identification [29-31,50] . DNA microarray is constructed in parallel manner via various methods including inkjet, pinmicroarray and lithography. Since millions of sequences are synthetized by single run with lengths comparable to conventional synthesis, such high-density microarrays enables high-throughput screening of target molecules.

Recent development in DNA microarray technology has emerged as a powerful alternative artificial DNA source by detaching synthesized molecules from the substrate. Cost can be reduced by approximately two orders of magnitude.

Column-based DNA synthesis	Microarray DNA
10,000 different oligonucleotide (200bp)	10,000 Microarray oligo pool (200bp)
$800,000 (0.4$/bp x 2,000,000bp) (from IDT, ultramer synthesis)	$2,500(from Customarray)

Table 1.1 DNA synthesis cost and throughput comparison table.

1.2 Separation of complex DNA pool

In spite of many advantages over conventional column based DNA synthesis, mixed pool microarray DNA needs to be separated in each contents except special applications such as pool library assay or parallel assembly. The most common way to separate complex mixture of DNA pool is vector cloning.

Vector cloning

Pure DNA samples from complex DNA pool such as large genomes or microarray DNA can be obtained by recombinant DNA technology. The essence of

recombinant DNA technology is the preparation of large numbers of identical DNA molecules. A DNA fragment of interest is linked to vector DNA molecules, which can replicate when introduced into a host cell. The inserted DNA is reproduced along with the vector. There are two types of vector; *E.coli* plasmid and bacteriophage λ vector. Here I briefly introduce general vector cloning process [38].

Conventional vector cloning process starts with pool amplification. After gel purification of amplified contents, DNA fragments are inserted into plasmid vector followed by bacterial transformation. The inserted DNA is replicated along with the rest of the plasmid DNA and segregates to daughter cells as the colony grows. Then cells are cultured on antibiotic selective plates by manual spreading and overnight incubation. Serial process of picking colonies which contain identical DNA and culture in Luria-Bertani media for plasmid extraction takes another several days depending on the number of colonies to pick. Finally experimenter chooses insertion containing plasmid through colony PCR and conducts plasmid extraction for Sanger derived identification. Since most of cloning fully relies on randomized manual selection of colony followed by individual capillary sequencing, the throughput is very low and unnecessary additional work is required.

DNA barcode

DNA barcode is a short DNA sequence that uniquely identifies a certain linked feature such as a gene or mutation. Since all short DNA sequence pair can be used for primer to amplify specific sequences, it is also applicable for sub-pool separation from complex DNA pool. In 2010 at nature biotechnology, Sriram Kosuri demonstrated scalable gene synthesis by selective amplification of DNA pools from microchips [42]. By using orthogonal primer set, they eliminates cross hybridization therefore selectively separate sub-pool containing precursor DNA sequences for each gene synthesis. However, approach requires quite large number of individual primer sets which turned into potential cost barrier and carries error from microarray DNA synthesis process.

1.3 Identification of DNA

Sanger sequencing

Sanger sequencing developed by Fred Sanger in 1975, was one of the most important art of Human Genome Project in 1990. This is a method of DNA sequencing based on the selective incorporation of chain-terminating dideoxynucleotides by DNA polymerase during in vitro DNA replication. The template DNA is copied repeatedly. The copying reactions stop when modified nucleotides called ‘chain terminators’ are added. By generating a large number of

fragments of different lengths that end in fluorescently labeled bases. Original sequence can be read by capillary gel electrophoresis which sorts those fragments by size to read out final fluorescent base through laser excitation. Generally Sanger method suffers from poor quality in the first tens of bases of the original sequence and low throughput due to one-by-one capillary electrophoresis process.

Pyrosequencing

Pyrosequencing is a method of DNA sequencing based on the 'sequencing by synthesis' principle relying on the detection of pyrophosphate release on nucleotide incorporation. Reactive solutions of A, C, G and T nucleotides are sequentially supplied to and removed from template DNA attached on the solid substrate. DNA polymerase incorporates the complementary dNTPs onto the template. In this step pyrophosphate (PPI) is or are released equivalently. Then ATP sulfurylase convets PPI to ATP that activates luciferase-mediated conversion of luciferin to oxyluciferin and generating visible light. This process is repeated with each of four letters until the DNA sequences of the single stranded template is determined [54].

The high demand for low-cost sequencing has driven the development of high-throughput next generation sequencing technologies that parallelize the

sequencing process, producing millions of sequencing information concurrently. Among various approaches, I would like to introduce Roche 454 Life Sciences technology that is a parallelized version of conventional pyrosequencing. The process starts with fragmentation of long target DNA sequence. After adaptor ligation, fragments are amplified by bead containing emulsion PCR to conduct single strand separation and amplification. Beads were denatured and put one bead into a well on a sequencing plate (selectively etched fiber bundle) with enzyme, ppiase and packing beads. Pyrosequencing is conducted in each separate well with parallel manner which dramatically increase the throughput upto one gigabase in a single day and lower down the cost compared with the Sanger sequencing.

1.4 Deterministic clone targeting

Conventional randomized colony separation and individual identification suffers from high demands of resources such as labor, cost and time. Such property is not appropriate for utilizing individual contents of high complexity microarray oligonucleotide pool. To overcome drawbacks of conventional approaches, I develop clone tracking algorithm on next generation sequencing plate and non-contact high-throughput pulse laser separation system. This ‘Sniper Cloning : massively parallel identification followed by deterministic separation’ directly utilized powerful NGS reading capacity to writing DNA ability and plays an import

role on synthetic biology in general.

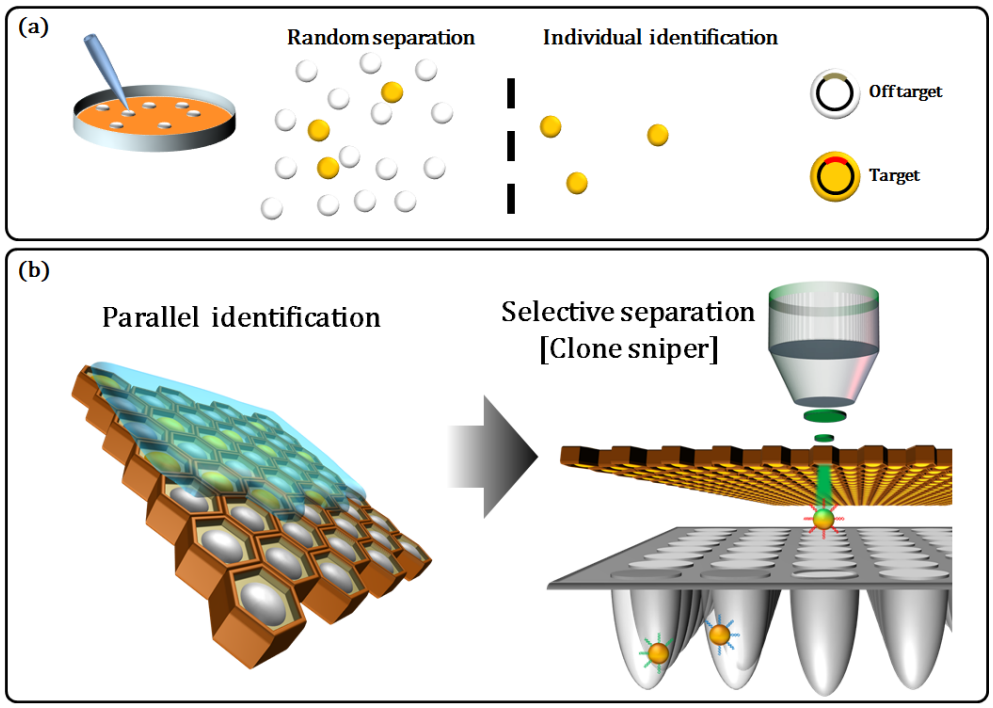


Figure 1.1 Comparison on conventional randomized cloning selection and Sniper Cloning method. (a) Convention cloning selection. Process consists of *in-vivo* molecular cloning, randomized colony picking, and individual capillary sequencing. (b) Sniper Cloning method. Massive molecular clones are prescreened using next generation sequencing platform. Non-contact and high-throughput separation system selectively separates target clones.

Chapter 2

Encoded microparticle based separation technique

The work presented in this section of the thesis involves microparticle based complex pool separation techniques. The microparticles which have distinctive IDs and probe oligonucleotides on their surfaces plays a floating microfilter that selectively separates the target single strand DNA from complex pool according to probe sequence. The fabrication method of color barcoded microparticles and magnetochromatic sphere are described.

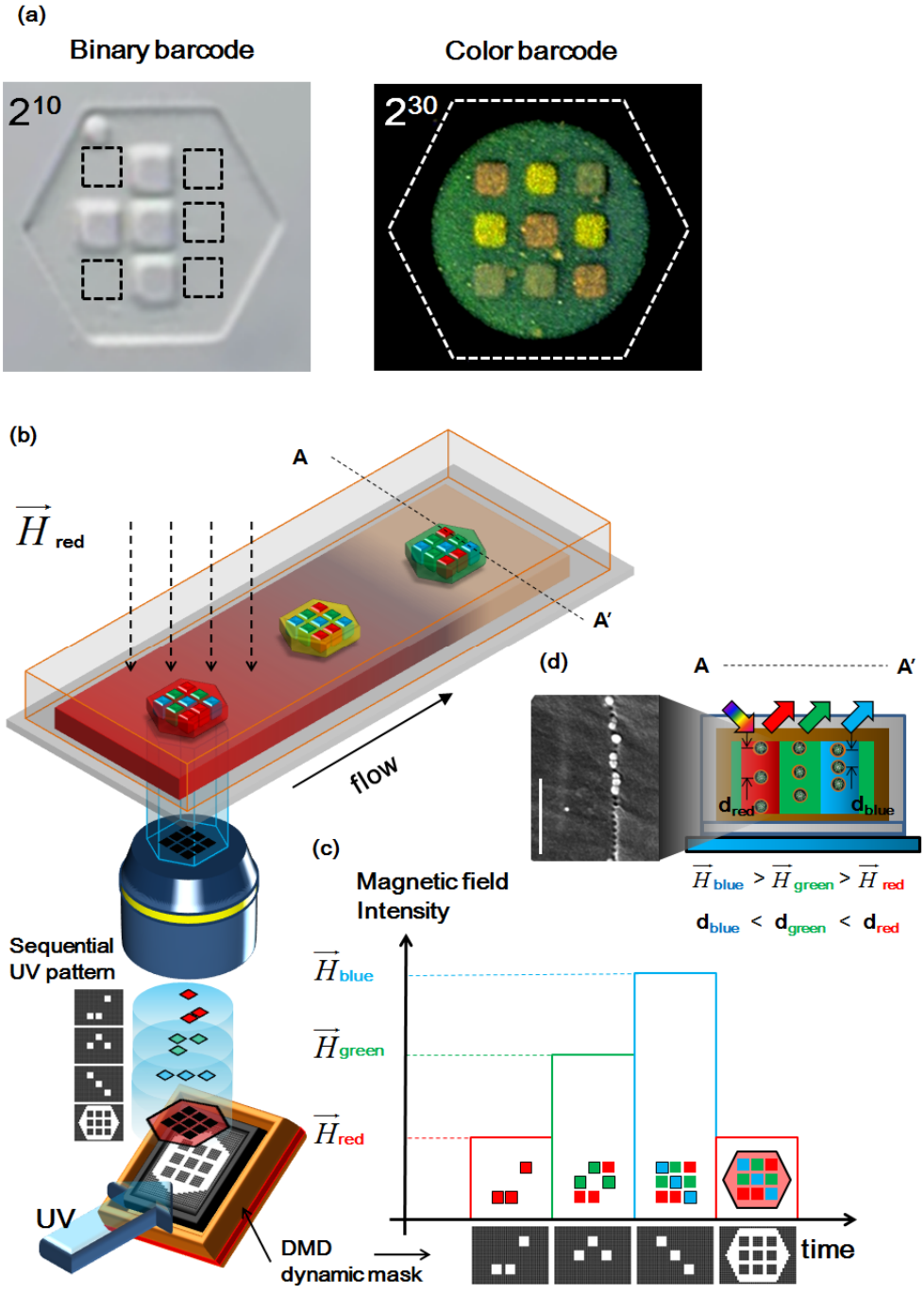

(a)
Binary barcode
Color barcode
2^{10}
2^{30}
(b)
A
A′
$\overrightarrow{H}_{red}$
flow
Sequential
UV pattern
UV
DMD
dynamic mask
(c)
Magnetic field
Intensity
$\overrightarrow{H}_{blue}$
$\overrightarrow{H}_{green}$
$\overrightarrow{H}_{red}$
time
(d)
A
A′
d_{red}
d_{blue}
$\overrightarrow{H}_{blue} > \overrightarrow{H}_{green} > \overrightarrow{H}_{red}$
$d_{blue} < d_{green} < d_{red}$

Figure 2.1 Single material (M-ink) based color barcoded magnetic microparticles. (a) Coding capacity comparison between a conventional binary barcode and a color barcode. (b) Conceptual description of the process of generating color barcoded magnetic microparticles. The system is composed of sequential processes involving the cooperative action of magnetic field modulation and spatiotemporally controlled UV exposure. (c) Time sequential modulation of the magnetic field according to a specific color and UV mask pattern for making color barcoded magnetic microparticle. (d) Cross-section of the polydimethylsiloxane (PDMS) microfluidic channel. By applying the external magnetic field, self-assembly of the magnetic colloids in the form of chain structures was used to create a fully reversible 1D photonic crystal as shown in the SEM image. A stronger magnetic field results in a shorter inter-particle distance, which corresponds to a shorter diffracted wavelength. Generated particles remain unbound from the channel wall due to an oxygen inhibition layer near the channel walls. Scale bars indicate 1 μm in (d). [58]

2.1 Color barcoded magnetic microparticle

Encoded particles have a demonstrated value for multiplexed high-throughput bioassays such as drug discovery and clinical diagnostics [1-2]. In diverse samples, the ability to use a large number of distinct IDs to assay particles is important to increase throughput [3]. Also proper handling schemes are needed to readout IDs on free floating probe microparticles. Here I create vivid, free floating structural color particles with multi-axis rotational control using a color tunable magnetic material and novel printing method [4]. Color barcoded magnetic particles offer a coding capacity easily into the billions with distinct magnetic handling capabilities including active stirring for improved reaction kinetics and active positioning for ID readouts in microscale environments [5]. A DNA hybridization assay has been done using the color barcoded magnetic microparticles to demonstrate multiplexing capabilities

Various encoding mechanisms have been proposed including spectroscopic [6-11] and graphical encoding to provide a sufficient number of IDs for microparticles [12-19]. Spectroscopic encoding, the most well established encoding scheme, commonly relies on both the color and intensity of light emitted from more than one coloring material such as fluorescence dyes [7,9] or quantum dots [8,10-11], which are embedded inside or attached on the surface of the microparticle. The overall spectrum of the encoded microparticles serves as an ID, which can be

readout by a spectrometer. However, the possible number of codes is practically limited due to the availability of the different indicating materials with minimal spectral overlaps. Moreover, the fabrication process is delicate since many different indicating materials must be precisely loaded into the microparticle with an exact stoichiometry. Graphical encoding is different approach that is based on the patterning of optically detectable elements on the microparticle surface. Binary barcoding is a popular coding scheme that has relatively large encoding capacity [12,14,17]. However, the number of optically resolvable features in limited by particle size, which places limits on the maximum graphical encoding capacity of a given particle.

Here I combine the easy to read nature of spectral encoding with greater coding capacity of graphical coding, improving on the best aspect of either approach. A typical example of color barcoded microparticles with 2D color barcodes is shown in Figure 2.1-(a). Compared with conventional binary barcoded particles, the use of color barcoding allows for a dramatic increase in encoding capacity. In this example, I have used 8 easily distinguishable colors to implement multilevel coding which provides a geometric increase in coding capacity. Practically speaking, consider that with 10 distinct bits in binary barcoding it is only possible to have 2^{10} ($\approx 10^3$) different codes. Here using a 10 bit system I demonstrate the capacity to create microparticles with 8^{10} ($= 2^{30} \approx 10^9$) easily readable color codes. Equally impressive

is that these 10 bit particles with up to 8 different color encodings can be created in approximately 1 second. This huge coding capacity of the color barcoded particles enables the encoding of much larger molecular libraries with smaller particle sizes.

2.1.1 Fabrication of color barcoded magnetic microparticle

Up until now the expression of multiple colors in microscale objects required many cycles of consecutive patterning of multiple color materials at high resolution, which is not practically scalable. I overcame this limitation by fabricating microparticles using a single M-Ink material. The process works through the repetitive tuning and fixing of structural color in M-Ink, which greatly simplifies the manufacturing process. Figure 2.1-(b) is a schematic diagram of multicolor barcoded microparticle generation. Instead of patterning multiple coloring materials, a sequential process is used involving cooperative actions of magnetic field modulation and spatially controlled UV exposure (Figure 2.1-(c)). First, a Polydimethylsiloxane (PDMS) channel is filled with M-Ink. The color of M-ink is then tuned by modulating an external magnetic field. The magnetic field changes the periodicity of 1D chain structures of the superparamagnetic colloidal nanocrystal clusters (CNCs) and the way these clusters reflect the corresponding wavelength of light [20-21]. Once the desired color for a specific barcode is set by the external magnetic field, I illuminate the UV mask pattern of the specific color barcode to

locally fix the color by ‘freezing’ the 1D chain structure of the superparamagnetic CNCs in M-Ink (Figure 2.1-(d)). Instead of using physical masks for the barcode pattern, a digital micromirror device (DMD) is used which acts as a computer controlled spatial light modulator. Making the next barcode bit using a different color is done by simply changing the pattern of DMD and the magnetic field intensity. This color tuning and fixing process takes approximately tenths of a second for each bit. Note that there is no need for realignment of the DMD or the fabrication area after each new color is added. . In addition, by virtue of the oxygen inhibition layer inside the PDMS channel, generated microparticles can move along the flow stream without being stuck to the channel walls, allowing for rapid successive microparticle creation using a microfluidic channel environment [23].

To monitor chain formation of superparamagnetic CNCs, I implemented a measurement setup as shown in Figure 2.2. Two electromagnets exert a magnetic field in the same direction to generate a perpendicular magnetic field with respect to the sample. The superparamagnetic CNCs/PEG-DA mixture was loaded between two coverslips with a 40μm-PDMS spacer. Transmitted light is sent to spectrometer through optical fiber and analyzed. Figure 2.3 shows the change of spectral transmittance under the on/off operation of the electromagnets. The sample starts to reflect light with a peak wavelength of 512nm under the constant magnetic field, and intensity of transmitted light decreases as a result. Approximately 70% of the

transition was accomplished within 150ms. The sample goes back to its original state after turning off the electromagnets.

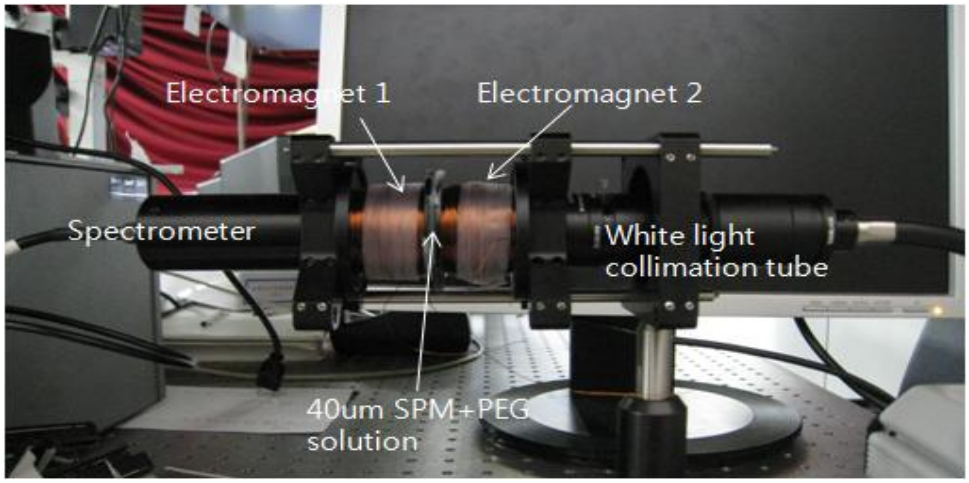


Figure 2.2 Step response measurement setup. The system is composed of two coreless electromagnets, a fiber-coupled spectrometer and a collimated white light source. A thin layer of superparamagnetic CNCs/PEG-DA mixture is located between two electromagnets. [58]

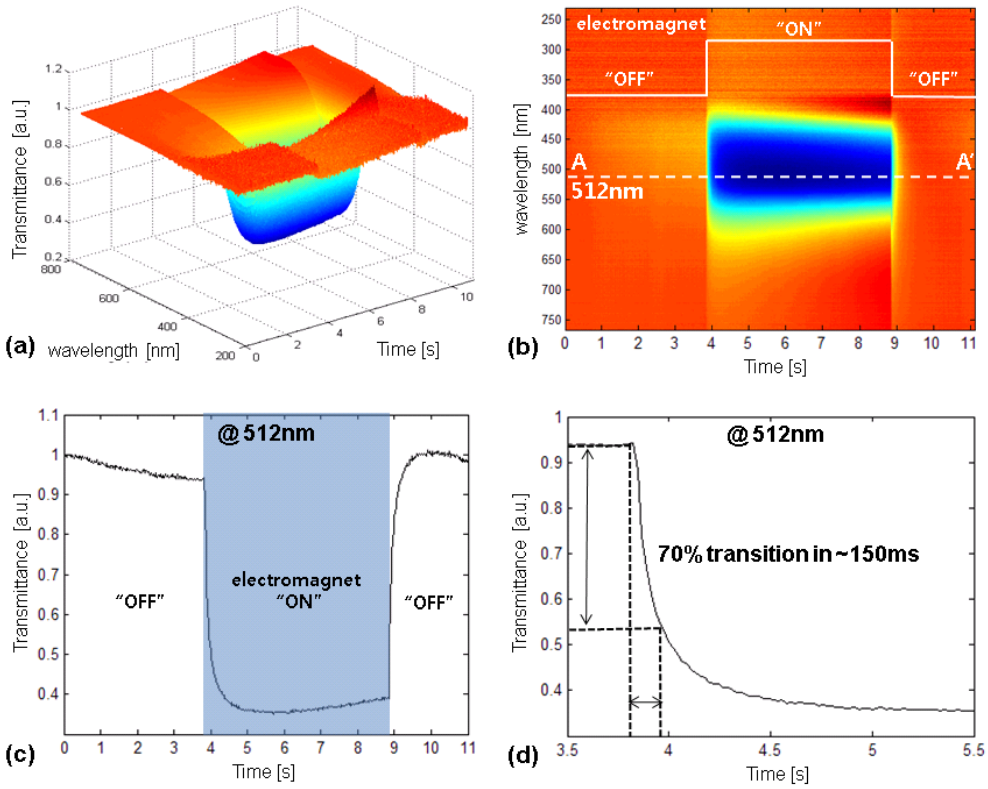


Figure 2.3 Step response of spectral transmittance of superparamagnetic CNCs/PEG-DA mixture. Spectral transmittance was determined by dividing the spectrum of transmitted light by the spectrum of the light source. (a) Data set for a single on/off operation of the electromagnet. Each spectrum was acquired every 22ms. (b) Top-view of (a). On/off operation of the electromagnet is clearly shown. (c) Change of transmittance with respect to time at a wavelength of 512nm. (d) Close-up view of (c) during the transition period. [58]

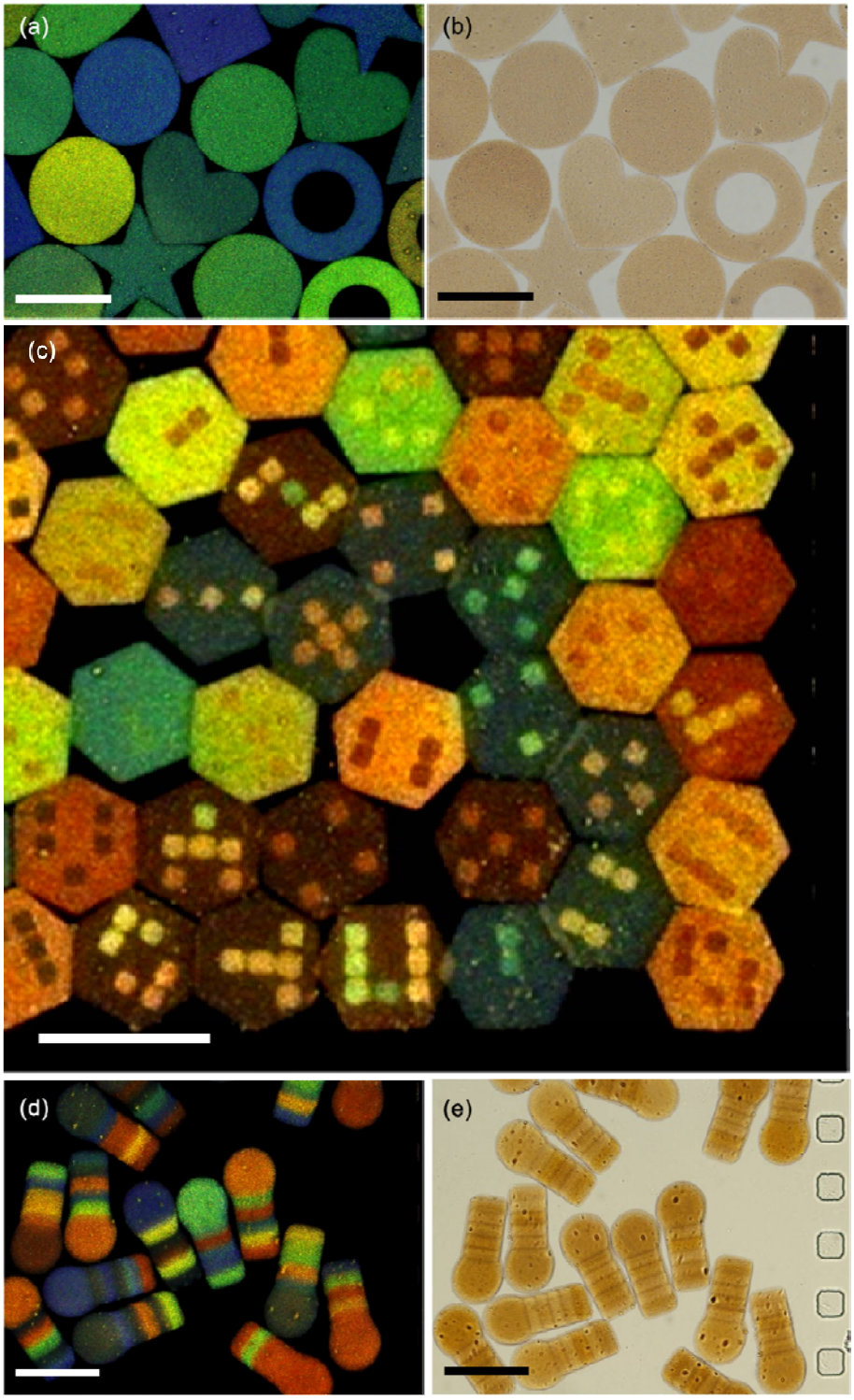

(a)
(b)
(c)
(d)
(e)

Figure 2.4 (a) Various shaped and colored particles and their tramsmission micrograph (b). (c) Hexagon type 2D color barcoded microparticles. (d) Bar-type 1D color barcoded microparticles and their transmission micrograph (e). Scale bars indicate 200 μm in (a,b), 500 μm in (c) and 250 μm in (d,e). [58]

.

Various types of color barcoded microparticles are demonstrated here for different coding and detection schemes, channel designs, and probe types (Figure 2.4). A hexagonal 2D barcoded microparticle (Figure 2.4-(c)) is suitable for close 2D packing in the microchannel while 1D color barcoded microparticles (Figure 2.4-(d)) are good for flow through detection methods. I created particles with up to 10 spatially separated barcodes with 8 distinct colors based on 7 structural colors and the intrinsic color of M-Ink. Though I conservatively chose only 8 different colors for coding with consideration of the reflection spectrum of M-Ink, optimization of the concept would result in a theoretically unlimited coding capacity. A transmission micrograph of the particles shows the intrinsic brownish color of the magnetite (Figure 2.4-(b),(e)). The difference between the colorful reflective image and the brownish transmission image indicates that the colors are not from pigmentation but from the structure of the superparamagnetic CNCs. Since the spectrum of generated structural colors has single peak value, a cheap charge coupled device (CCD)

camera can be used to decode IDs from microparticles using simple RGB level data analysis.

2.1.2 Magnetic handling for multistep reaction

Another important aspect of coded particle based assays, other than the encoding capacity, is the handling scheme of the particles in suspension. It is not trivial to separate particles from solution mixtures. Up until now, magnetic handling has been widely used for active separation of desired contents from the carrier solution. By applying an external magnetic field to a solution, a mixture containing magnetic beads as well as the molecules bound to those magnetic beads can be selectively sorted out from the solution mixture. Therefore it is beneficial to incorporate magnetic materials into the particles for assays involving solution exchange [24]

Under the assumption that a single particle is a point-like dipole, the magnetic force acting on the particle with volume V_p, magnetic susceptibility of particle χ_p and magnetic susceptibility of fluid χ_f can be expressed as

$$\vec{F}_m = \frac{V_p(\chi_p - \chi_f)(\vec{B}\cdot\nabla)\vec{B}}{\mu_0}$$

(μ_0 : permeability of free space). This equation should be modified when the magnetization of the particle gets saturated, but the particles of interest were usually manipulated under the magnetic field of < 2kOe in my case. Hydrodynamic drag

force acting on the particle by the fluid with viscosity η in low Reynolds number regime is

$$\vec{F}_d = 6 \cdot \pi \cdot \eta \cdot R_{p\_eq} \cdot \xi \cdot \vec{v} \,,$$

where ξ is dynamic shape factor and R_{p_eq} is equivalent volume radius of spherical particle having the same volume as a non-spherical particle. I used the dynamic shape factor 1.28 for a thin disk with length 30μm and diameter 250μm (aspect ratio of 0.12), which is similar to geometrical factor of color barcoded magnetic microparticles. The effect of Brownian motion and gravitational and buoyant forces can be neglected due to the large particle size and the actuation direction, which is perpendicular to gravitational force. I also assumed that interaction of dispersed particles can be neglected due to the small fractional volume of the particles in solution.

By combining the two formulas above, I can get the analytic solution for terminal velocity of the particle to be

$$\vec{v} = \frac{R_{p\_eq}{}^2 \cdot (\chi_p - \chi_f) \cdot \nabla \vec{B}^2}{9 \cdot \eta \cdot \mu_0 \cdot \xi} \,,$$

with relaxation time $\tau = \frac{2\rho R_{p\_eq}{}^2}{9\eta\xi}$. The relaxation time is short compared to manipulation frequency as shown in Figure 2.5-(b), so I only consider the terminal velocity in this case. Figure 2.5-(c) shows the dependency of particle velocity on the

radius of the particle at a distance of 1cm from the magnet. For the fluid I chose ethanol and used the magnetic field profile shown in Figure 2.5-(a).

In the case of a spherical particle with the radius of 71μm, which is an equivalent volume radius of color barcoded magnetic microparticles, and with the distance of 1cm from the magnet, the terminal separation speed $\left|\vec{v}_{th}\right|$ of the particle is ~ 4mm/s and increases as the particle gets closer to the magnet due to the increase in magnetic field gradient and magnetization of the particle. The experimental result shows consistency with this expectation.

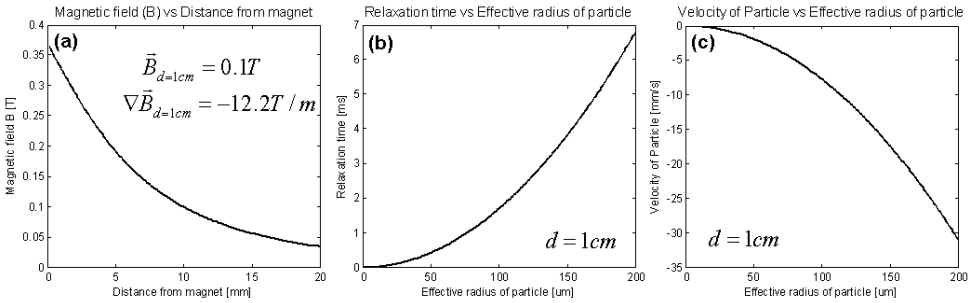
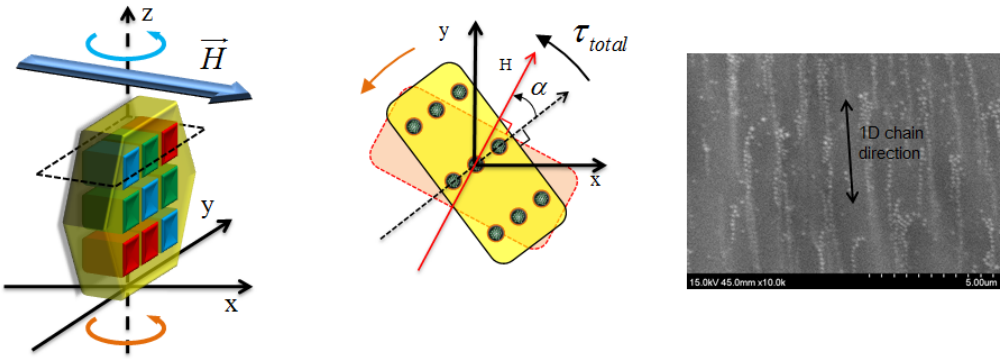


Figure 2.5 Theoretical estimation of separation characteristics of color barcoded magnetic microparticles. (a) Magnetic field profile with respect to the distance from the magnet. (b) Relaxation time according to the size of the particle at the distance of 1cm from the magnet. (c) Velocity of the particle dependent on the radius of the particle at the distance of 1cm from the magnet. [58]



(a) (b) (c)

Figure 2.6 A unique magnetic manipulation technique of structural color encoded magnetic microparticles for multi-step bio-chemical assays using magnetic axis control. **(a)** A schematic diagram of color barcoded magnetic microparticle rotation. **(b)** The total magnetic torque is simply N times the torque on a single chain and maximum value is obtained as $N\frac{4\mu_0 n^2 \chi^2 R^3 \pi}{3d^3}H^2$ when α is 45°. **(c)** SEM image of 1D chain structure located in the cross section of the generated particle. [58]

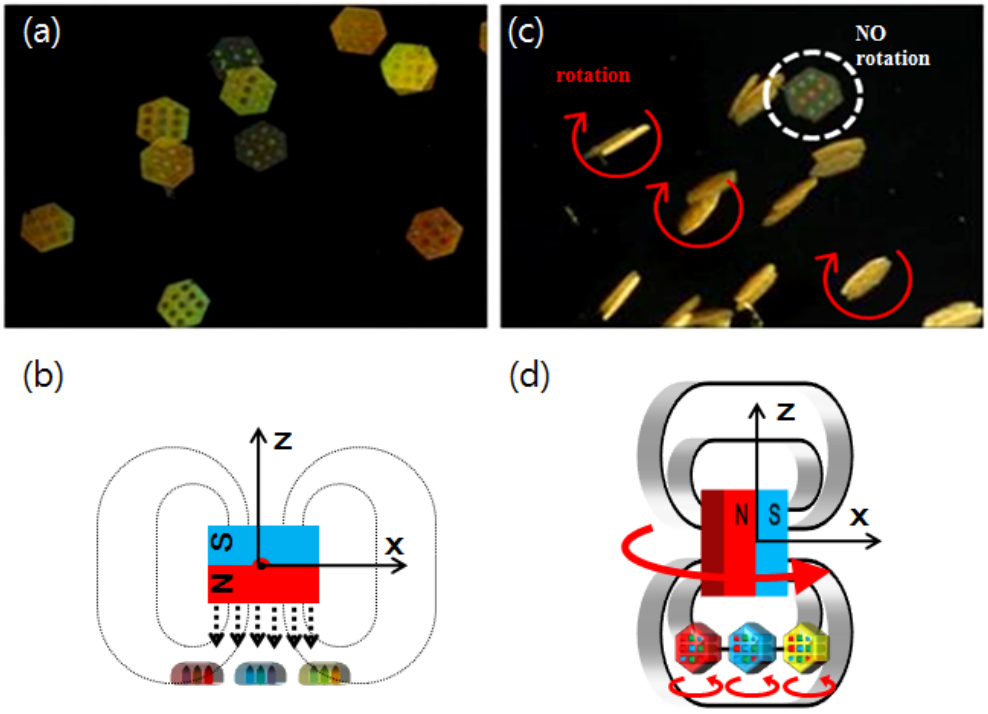


Figure 2.7 (a) Reflection microscopy of encoded particles under vertical magnetic field lines. The barcoded particles are displaying IDs on 2D surface of the vial. (b) External magnetic field line when 2D reading mode. (c) Rotating encoded particles (Movie S3). The non-rotating particle, which has no 1D chain CNCs structures, shows the brownish intrinsic color of the magnetite. (d) External magnetic field line when 3D reaction mode. [58]

One significant advantage of color barcoded magnetic microparticles over conventional magnetic beads is the capability of multi-axis manipulation allowing for flipping, spinning, and rotation. Magnetic nanoparticles located inside the microparticles form a 1D chain structure (Figure 2.6-(b),(c)). As a result, particles are not only magnetically separable but have also been engineered to align with external magnetic field lines (Figure 2.6-(a)) [25-26]. Using this property, color coded particles themselves are capable of working in tandem as millions of microscale rotating stirrers. This capability is impossible with conventional magnetic beads, which contain randomly distributed magnetic materials. As shown in Figure 2.7, a rotating horizontal external magnetic field rotates vertically aligned encoded particles around its own vertical axis, allowing for rapid mixing and thus maximizing the reaction between the probe and target in solution. Note that one can easily control the magnetic anisotropy of the particles during the UV exposure step by simply controlling the direction of the external magnetic field. The movie clip of this rotational action is included in the online supporting material. One non-rotating barcoded particle in Figure 2.7-(c) was intentionally fabricated with randomly distributed magnetic nanoparticles, while all other rotating barcoded particles were fabricated with aligned 1D nanoparticle chains. This clearly shows that the 1D chain structure of the magnetic nanoparticles is necessary for rotation.

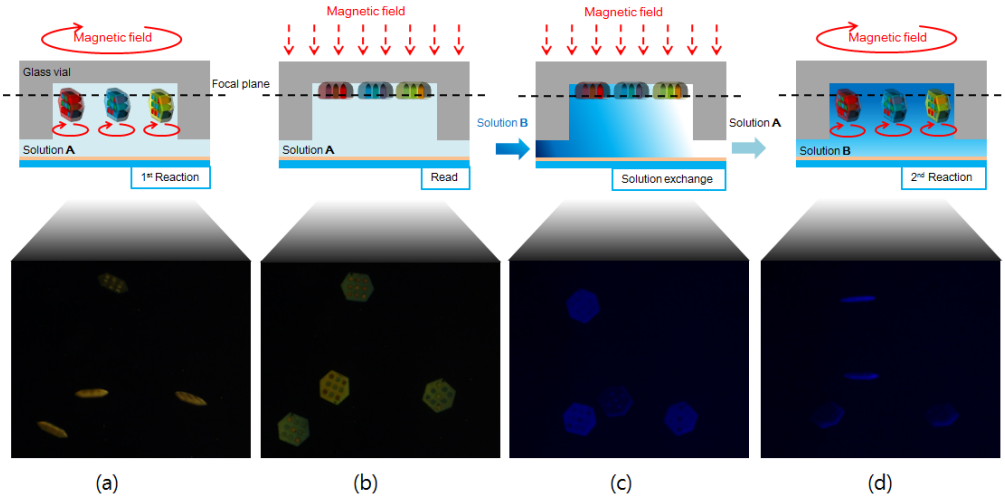


Figure 2.8 (a) to (d) 3D reaction and 2D reading scheme for a multistep reaction. (a) First 3D reaction, (b) 2D reading, (c) solution exchange and (d) second 3D reaction in solution B. [58]

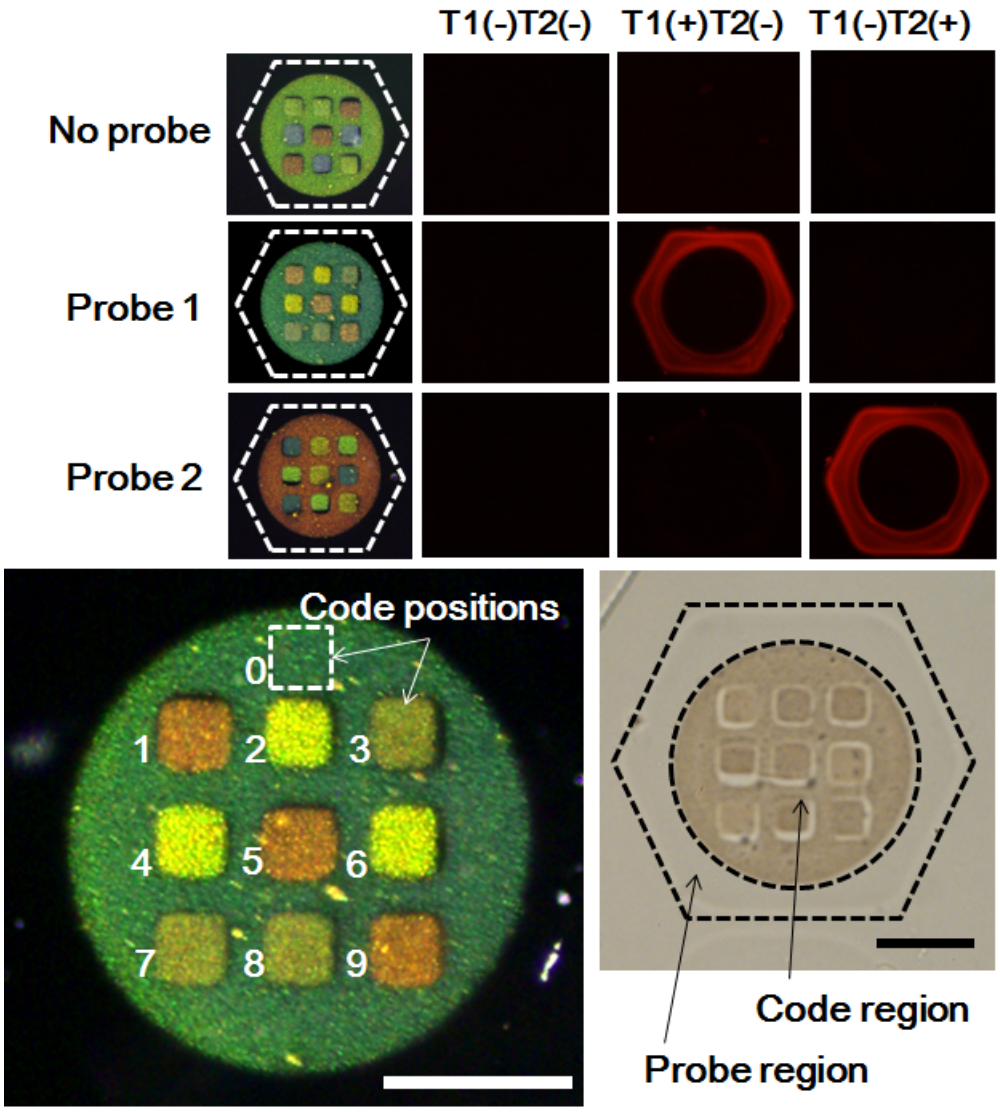
Based on the magnetic handling capabilities of color barcoded particles, I have demonstrated a multistep assay that consists of an in-situ 3D reaction and 2D reading process (Figure 2.8) The idea is first to perform a reaction in solution assisted by active stirring and then to read out the IDs of the particles by pulling them onto the surface of a vial. A multi-axis permanent magnet rotation system has been developed to magnetically manipulate the particles. Depending on the direction of the external magnetic field lines, the color barcoded magnetic microparticles can be laid out or stood upright with respect to the vial surface. After setting the

magnetic field horizontally to align the particles perpendicular to the vial surface the multi-axis magnet rotational device was rotated on the Z axis by a motor to activate rotational stirring of the color barcoded magnetic microparticles. Using magnetic apparatus the probe molecules located on the microparticles have a greater chance to bind with target molecules via enhanced mass transportation (Figure 2.9). After the reaction step, the color barcoded particles are laid down flat by aligning the external magnetic field vertically to the vial surface, so that their IDs can be read using a 2D imager. Due the ability to control flow resistance by holding the microparticles in place using simple magnetic control, this process can be used for repeated bioassays involving multi-step solution exchange.

The rotational speed of the particle is dominated by solution viscosity and magnetic field strengths. In typical solutions I found it possible to rotate the particles at a rotational frequency of several hundred RPM. The total torque of the microparticle having 1D chains of superparamagnetic nanoparticles under the external magnetic field is simply N times the torque on single chain due to the bilateral symmetry of the structure. The equation below shows the total torque on a microparticle in detail

$$\tau_{total} = N\tau_i = N\frac{3\mu_0 m^2}{4\pi}\sin(2\alpha)\frac{n^2}{d^3} = N\frac{4\mu_0 n^2 \chi^2 R^6 \pi}{3d^3}H^2\sin(2\alpha)$$

where N is the number of chains inside a single microparticle, n is the number of superparamagnetic CNCs in a single chain, R is the radius of a superparamagnetic CNC and χ is the initial mass susceptibility of superparamagnetic CNCs. Here, the total magnetic torque is simply N times the torque on a single chain since inter-chain interactions are ignored in the microparticles. This result implies that the motion of microparticles is easily controlled by changing the structural geometry or inter-particle distance in chain structures. The maximum total magnetic torque is obtained as $N\frac{4\mu_0 n^2\chi^2 R^3\pi}{3d^3}H^2$ when α is 45°.



Pos.	(R,G,B)	CODE	Pos.	(R,G,B)	CODE
0	(50 , 114, 60)	(1,2,3)	5	(173, 121, 36)	(3,2,2)
1	(180, 124, 38)	(3,2,2)	6	(206, 235, 7)	(4,4,1)
2	(210, 236, 8)	(4,4,1)	7	(147, 153, 45)	(3,3,2)
3	(134, 142, 41)	(3,3,2)	8	(145, 152, 46)	(3,3,2)
4	(209, 235, 8)	(4,4,1)	9	(172, 119, 34)	(3,2,2)

Figure 2.9 Multiplexed biomolecule assay using M-ink probes. **(a)** Reflective micrographs of three different encoded particles with different DNA oligomer probes and their multiplexed analysis based on DNA oligomer target hybridization. Each M-ink probe consists of a code region and probe region as shown in **(d)**. The probe region is loaded with 12.5μM DNA oligomer probe and shows a fluorescence signal after incubation with 1μM fluorescence-labeled complementary DNA oligomer target. All fluorescence images are obtained under the same imaging condition. **(b)** Fluorescence intensities from M-ink probes (Probe 2, 12.5μM) after incubation with complementary target DNA oligomer (T2, 100nM). The probe particles were incubated in target-containing solution between two PDMS-coated glass slides with a 2mm-thick PDMS spacer. I induced reaction enhancement by rotating the microparticles during incubation. The reaction enhancement occurred on both the body and the edge of the rotating particles. The fluorescence images of three representative particles obtained under the same imaging conditions are also shown. **(c)** Magnified view of a single encoded particle with probe 1 after hybridization with its complementary target oligomer (T1). Each label in the figure indicates the position of each color bit in code region. **(d)** Bright field micrograph of a representative particle. The dashed-lines are drawn to clearly indicate the boundary of each region. **(e)** Code reading of the particle in **(c)**. RGB levels from individual codes are read from a color image taken by true-color CCD camera. Ten individual

“code reads” are grouped and highlighted with different colors. All scale bars indicate 100μm. [58]

2.1.3 DNA separation

Finally, I demonstrated an example of a multiplexed biomolecule assay using the color barcoded particles for detection and identification of DNA hybridization. The coding capacity of this color barcoded microparticle can be in the billions with particle size smaller than a conventional probe spot in a DNA microarray (< 100 micrometer). The color barcoded particle consists of a code region and an oligo probe region as shown in Figure 2.9-(d). I separated the oligo probe region from the code region to avoid spectral overlap between the fluorescence signal for detection of hybridization and the structural color signal for color encoding. The code region is synthesized from M-Ink, and the probe region is formed from a mixture of PEG-DA and buffer solution with acrylate-modified DNA oligomer probes. DNA strand detection using the DNA-incorporated hydrogel is a simple approach with high sensitivity [17,27]. To demonstrate the specificity of the analysis, 12.5 μM DNA probes with different nucleotides sequences (Probe 1, 5'-ACA CTC TAC AAC TTC-3', Probe 2, 5'- ATC AGA TTG GTT AGT-3', and no DNA probe as a control) were incorporated into different color coded microparticles (Figure 2.9-(a)). DNA oligomer targets (1 μM) were then introduced and incubated for 10 minutes. Only the particles with DNA probe complementary to the DNA oligomer target show fluorescence. By rotating the particle during the incubation, I can increase DNA hybridization kinetics as shown in Figure 2.9-(b)

The decoding process is simple and compatible with ordinary microscopes and color imagers. Decoding of a single encoded particle for target molecule identification is shown in Figure 2.9-(c). RGB levels of individual code positions are obtained from reflective micrograph readings using a full color CCD. The resulting RGB values are digitized to specific code levels. I divide each RGB value into 4 levels, resulting in 4 possibilities for each R, G, and B value. Here, I chose the encoding and decoding scheme of a representative microparticle using square-shaped color bits with 10 code positions and 4 color variations. Again the capability for multiplexing is virtually limitless.

2.1.4 Summary

In summary, I have demonstrated fabrication and application of color barcoded magnetic microparticles based on structural colors from a magnetically color tunable material called M-Ink. In doing so, I have increased encoding capacity into the billions using brilliant, easily distinguishable color coding. The use of color tunable M-Ink also enabled cost-effective and scalable manufacturing of the color barcoded microparticles by eliminating need of multiple coloring materials. Based on the uniquely engineered magnetic anisotropy of the barcoded particles, distinctive magnetic manipulation schemes such as active stirring and ID readouts

are demonstrated. In order to show the feasibility of this method to do a practical multiplexed bioassay, DNA hybridization was performed using the color barcoded particles. I believe the color barcoded magnetic microparticles presented here greatly surpass previous limits in microparticle color encoding and will eventually offer biologists an affordable, superior alternative to bead based bioassays.

2.1.5 Materials and methods

Materials M-ink is composed of superparamagnetic colloidal nanocrystal clusters (CNCs) and photocurable resin monomer solution (S1). I used ethoxylated trimethylolpropane triacrylate (ETPTA, Sigma-Aldrich) as a photocurable resin with 10 wt% of 2,2-dimethoxy-2-phenylacetophenone (DMPA, Sigma-Aldrich) as a photoinitiator for the magnetic actuation demonstration. In the DNA hybridization assay demonstration, superparamagnetic CNCs dispersed in a monomer solution of 3:1 poly(ethylene glycol) diacrylate (PEG-DA, Sigma-Aldrich, Mn=700) : deionized water with 10 wt% DMPA were used for the coded region. Probe oligomer solutions for the DNA oligomer probe region were prepared in 3:1 PEG-DA:TE buffer (10mM Tris pH 8.0, 1mM EDTA) with 10 wt% DMPA and 12.5μM DNA oligomer probe (S3). Target oligomer solutions were prepared by adding 1μM Cy3-labelled DNA oligomer target in TE buffer with 0.2M NaCl (Daejung Chemicals) and 0.5% sodium dodecyl sulfate (SDS, Daejung Chemicals). I also used 3:1 PEG-DA:TE buffer as a wash buffer in every wash step.

Microfluidic channel fabrication Polydimethylsiloxane (PDMS) elastomer (Silgard 184, Dow Corning) was poured onto a silicon wafer which was patterned with SU-8 photoresist (SU-8 2015, MicroChem), and thermally cured for 15min on a 150˚C hot plate. The replica of the PDMS mold was bonded to the PDMS coated glass slide after oxygen plasma treatment with plasma cleaner

(CUTE-MP, Femto Science). For strong adhesion between the PDMS channel and PDMS coated glass slide, the bonded microfluidic device was heated on a 150˚C hot plate for 10min.

Magnetic field generation and maskless lithography Color barcoded magnetic microparticles were generated by photopolymerization using a spatial light modulator (S4). An optical microscope (IX71, Olympus), UV source (200W, mercury-xenon lamp, Hamamatsu) and digital mirror device (DMD, Texas Instrument) were aligned to expose a UV. An electromagnet was placed above the microfluidic device and calibrated using a gaussmeter (455 DSP Gaussmeter, Lakeshore).

2.2 Magnetochromatic microspheres

In multiplex assay system such as DNA pool separation, magnetic properties of probed microstructure are useful for multistep reaction, separation and small volume reaction. Here I introduce a new type of magnetically responsive photonic system whose orientation and consequently photonic property can be easily controlled by using external magnetic fields. Specifically, the fabrication of microspheres involves [57].

2.2.1 Fabrication of magnetochromatic microspheres

Magnetochromatic microspheres have been fabricated through instant assembly of SPM colloidal particles inside emulsion droplets of UV curable resin and then an immediate UV curing process to polymerize the droplets and fix the ordered structures. Unlike the previously reported “opal balls” whose orientation cannot be controlled, fixing of SPM chains makes microspheres magnetically “polarized” so that their orientation becomes fully tunable as the SPM chains always tend to align along the external field direction. Many copies of photonic crystal microspheres can be fabricated in a single process, and their orientation can be synchronically tuned to collectively display a uniform color. Compared to the magnetically tunable photonic crystals, the photonic microsphere system does not involve the nanoparticle assembly step and, therefore, has several advantages including long-term stability of optical response, improved tolerance to

environmental variances such as ionic strength and solvent hydrophobicity, and greater convenience for incorporation into many liquid or solid matrices without the need of complicated surface modification.

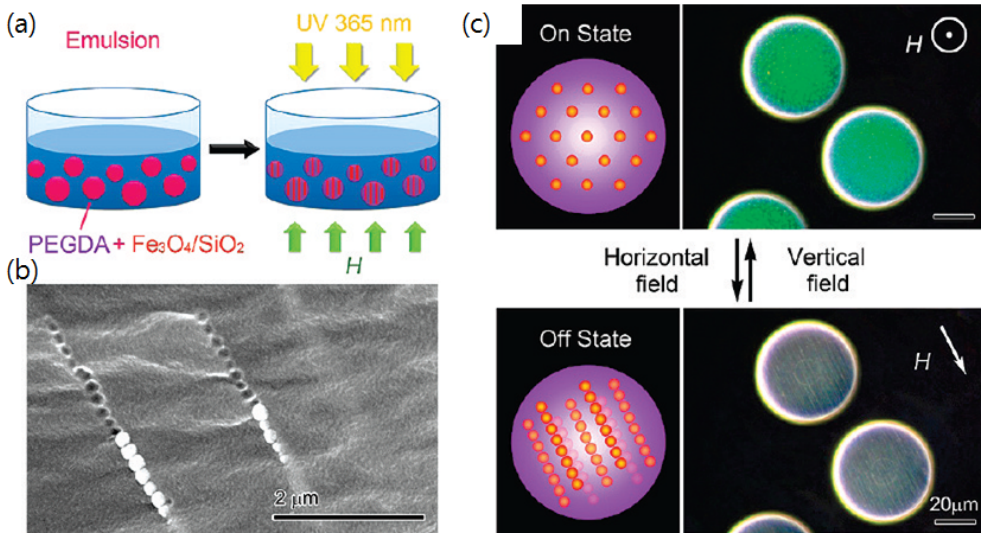


Figure 2.10 (a) UV initiated polymerization of the oligomers in emulsion droplets fixes the periodic structures inside the microspheres and retains the diffraction property. (b) SEM images of Fe3O4@SiO2 particle chains embedded in a PEGDA matrix. (c) Schematic illustrations and optical microscopy images for the magnetochromatic effect caused y rotating the chain-like photonic structures in magnetic fields. [57]

2.2.2 Optical response of magnetochromic microspheres

Depending on the direction of the external magnetic field, the particle chains may be suspended at any intermediate stage between the on and off states with a specific tilting angle (θ). I have investigated the dependence of diffraction peak wavelength (λ) and intensity on the tilting angle (θ) using an optical microscope coupled with a spectrometer (Figure 2.11). While the magnetic field is tuned within the plane constructed by the incident light and back scattered light, the diffraction from an isolated microsphere is recorded correspondingly by the spectrometer, as schematically shown in Figure 2.11-(a). It has been found that the diffraction peak blue shifts with decreasing intensity when the magnetic field direction is manipulated away from the angular bisector of incident light and back scattered light ($\theta \approx 14.5°$). Figure 2.11-(b) shows the spectra and corresponding microscopy images when the angle θ is tilted from +10° to -30°. Such a change in the diffraction peak position and intensity closely resembles the characteristics of an onedimensional Bragg photonic crystal, as proven by the close match between the experimental results and theoretical simulations.

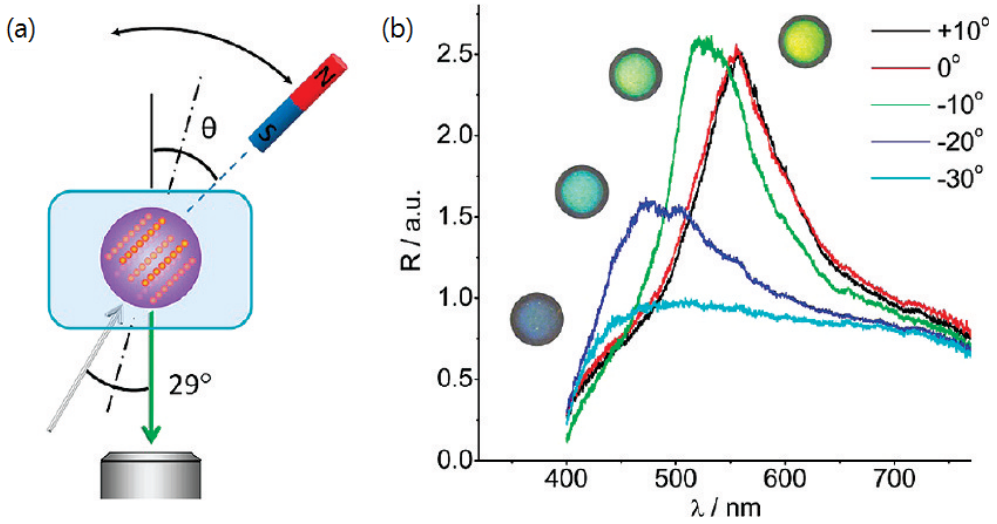
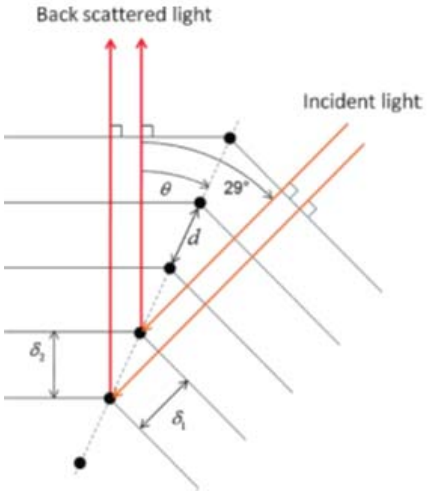


Figure 2.11 (a) Schematic illustrations of the experimental setup for studying the angular dependence of the diffraction property of the magnetochromatic microspheres. (b) Reflection spectrum and corresponding digital photo recorded from a single Fe3O4@SiO2/PEGDA microsphere at different tilting angles. [57]



$$\delta = \delta_1 + \delta_2 = nd\cos(29°-\theta) + nd\cos\theta \quad (1)$$

$$\phi = \phi_1 + \phi_2 = \frac{2\pi}{\lambda}\delta \quad (2)$$

$$A = f + fe^{i\phi} + fe^{i2\phi} + \ldots + fe^{in\phi} + \ldots$$
$$= f\sum(1 + e^{i\phi} + e^{i2\phi} + \ldots + e^{in\phi} + \ldots) = f\frac{1}{1 - e^{i\phi}} \quad (3)$$

$$|A|^2 = AA^* = f^2\left(\frac{1}{1 - e^{i\phi}}\right)\left(\frac{1}{1 - e^{-i\phi}}\right) = \frac{f^2}{2 - (e^{i\phi} + e^{-i\phi})}$$
$$= \frac{f^2}{2(1 - \cos\phi)} \quad (4)$$

$$\phi = \frac{2\pi}{\lambda}\delta = 2m\pi \quad (5)$$

$$\lambda = \delta = nd[\cos(29°-\theta) + \cos\theta] \quad (6)$$

Figure 2.12 Schematic illustration of geometric relation of incident light, back scattered light, and chain-like photonic structures for the simulation of diffraction wavelength at different tilting angles. [57]

The dependence of diffraction wavelength (λ) upon magnetic field tilting angle (θ) was simulated by formula (λ) nd[cos (29° - θ) + cos θ]) using MatLab. The refractive index (n) of PEGDA is 1.46, and the interparticle spacing within the chains is calculated to be 200 nm using the diffraction wavelength at Θ) 0°. I consider the chain-like structure as an ideal onedimensional photonic crystal composed of dots with negligible volume. According to the physical model in Figure 2.12, the optical path difference (δ) and phase difference (φ) can be expressed by eqs 1 and 2. Then, total amplitude (A) and intensity of back scattered light can be calculated by eqs 3 and 4. Since the maximum back scattering condition is only

achieved as the phase difference satisfying the eq 5, the final expression of diffraction wavelength will be deduced as eq 6. The simulated results are compared with the experimental data in Figure 2.13, both of which indicate that the diffraction peak wavelength has a maximum value when the particle chain is tilted along the angular bisector of incident light and back scattered light ($\theta \approx 14.5°$). When the tilting angle is manipulated away from 14.5°, which means increase from +20° to +50° or decrease from +10° to -30°, the diffraction peak blue-shifts and its intensity gradually decreases. This behavior is consistent with the typical characteristics of one-dimensional Bragg photonic crystal. Beyond -30° or +50°, the diffraction intensity is very low so that the photonic state of the microsphere can be practically considered as “off”.

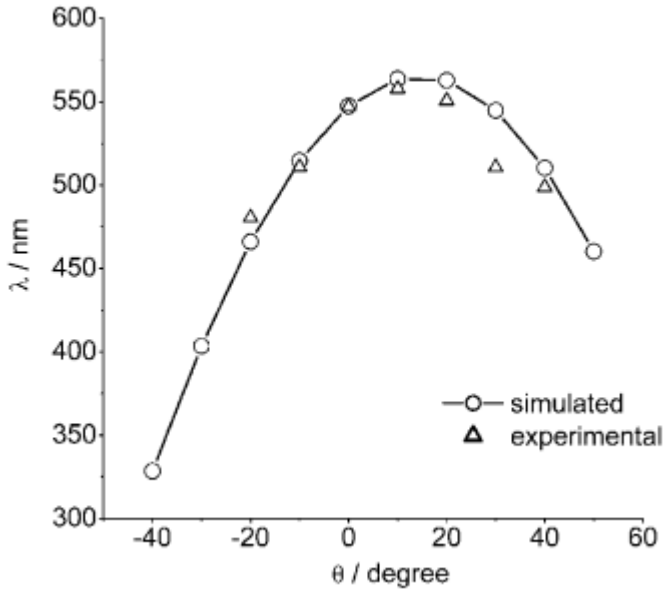


Figure 2.13 Simulation and experimental data showing the dependence of diffraction wavelength (λ) of a single Fe3O4@SiO2/PEGDA microsphere on the magnetic field tilting angle. [57]

The switching of diffraction could be accomplished rapidly(<1 s) in a sufficiently strong magnetic field. The turning frequency of the microspheres was measured with a test platform built with a halogen light source, a spectrometer, and a rotating magnet unit with a geared DC motor as shown in Figure 2.14. The rotating plate with NS and SN magnets standing alternately will produce a periodical vertical (1100-1200 G) and horizontal magnetic field (300-400 G), whose frequency can be simply controlled by the rotating speed of the plate. Figure 2.15 shows the diffraction of microspheres in a 1.22 and 3.33 Hz vertical/ horizontal alternating magnetic field, demonstrating that the photonic microspheres can be rotated quickly. It is noted that the rotating amplitude gradually decreases with the increase of turning frequency, primarily due to the relatively weak horizontal field strength. When the frequency is higher than 7 Hz the rotation of microspheres cannot catch up with the external field variation so that they seem to simply vibrate around the vertical state and the diffraction remains on all the time. The switching frequency might be further improved when the microspheres are dispersed in a less viscous solvent or tuned

in magnetic fields with higher strengths. Incorporation of photonic crystals into microspheres allows tuning of the photonic property by simply controlling the sphere orientation, making it very convenient to create bistable states that are required for many applications such as displays. Here I demonstrate a simple

switchable color display system in which the color information can be rewritten multiple times by means of a magnetic field. The basic idea is to create bistable states by embedding the microspheres into a matrix that can be switched between liquid and solid states. Many long chain hydrocarbons and short chain polymers, such as paraffin and poly (ethylene glycol), have melting points slightly above room temperature. When heated, the matrix material melts, allowing the display of colors by aligning the microspheres using magnetic fields. When the system is cooled to room temperature, the matrix solidifies and the orientation of microspheres is frozen so that the color information remains for a long time without the need of additional energy. An external magnetic field cannot alter their color once the orientation of microspheres is fixed by the matrix. Reheating the matrix materials, however, will erase the particular color by randomizing the orientation of the microspheres or by magnetically reorienting the microspheres to a completely “off” state.

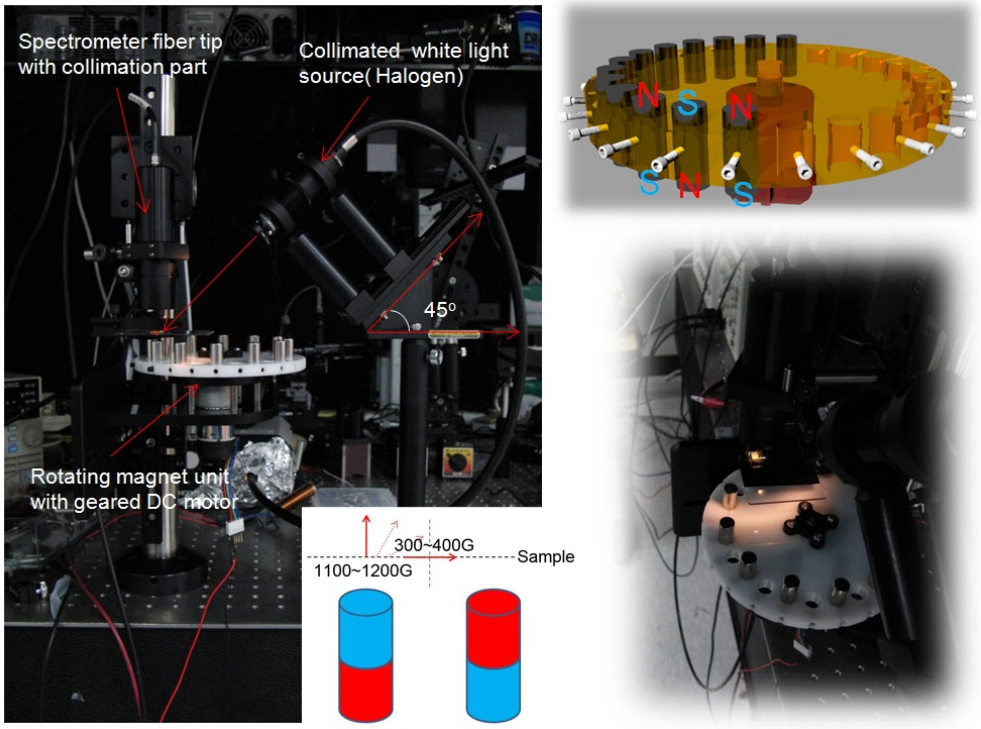


Figure 2.14 Digital photo images of the test platform for measuring the turning requency of microspheres in a periodical vertical/horizontal magnetic field. It includes a Halogen white light source, a spectrometer and a rotating plate with NS and SN magnet standing alternately. [57]

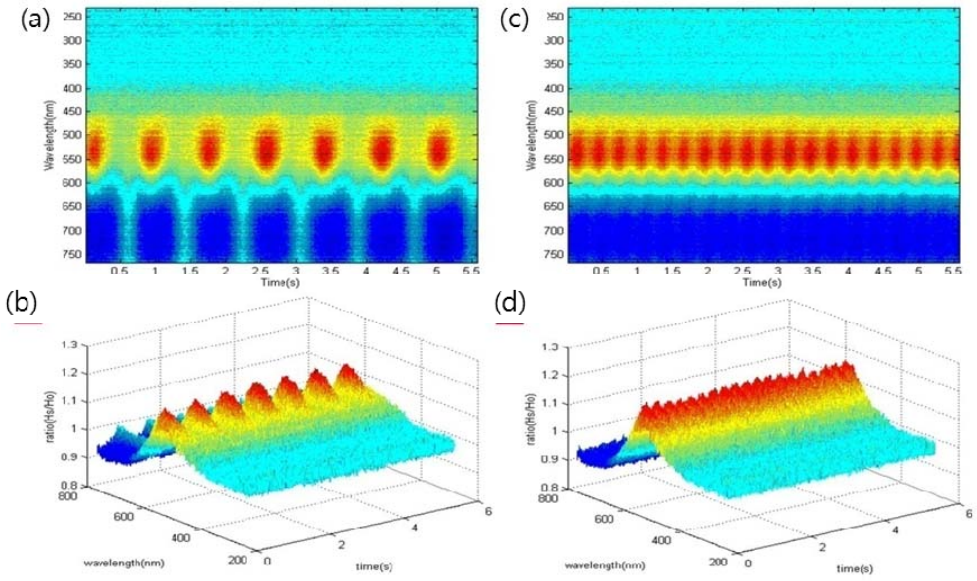


Figure 2.15 Optical response of Fe3O4@SiO2/PEGDA microspheres in a (a, b) 1.22 and (c, d) 3.33 Hz vertical/horizontal alternating magnetic field. Hs/Ho is the ratio of reflection with H field to that without H field. [57]

2.2.3 Summary

In summary, magnetochromatic microspheres have been prepared through a simultaneous magnetic assembly and UV curing process in an emulsion system. Superparamagnetic Fe3O4@SiO2 colloidal particles are self-organized into ordered structures inside emulsion droplets of UV curable resin, followed by an immediate UV curing process to polymerize the droplets and fix the ordered structures. By rotating the microspheres, it is very convenient to control the orientation of the magnetic chains and thereby the diffractive colors using external magnetic fields. Many copies of microspheres can be produced using the simple process, and can be tuned by external fields to collectively display uniform colors. Proposed fabrication method can reduce considerable amount of efforts to prepare bead based separation and its product enables simple magnetic manipulation in volume reaction.

Chapter 3

Complex pool separation technique based on pulse laser

The microstructures or microspots which carry probe oligonucleotide sequences filters out target sequence from a complex mixture pool. To utilize separated target molecules attached on the probe region, following physical separation process must be carried out. This chapter describes, the pulse laser driven microparticle and microarray separation technique by radiation pressure or substrate ablation. Firstly, the encoded magnetic microparticles having probe sequence on their surfaces are produced by flow lithography and subsequently separated by pulse laser radiation pressure after hybridization selection of complex pool. Secondly, the probe spot of microarray separation technique is demonstrated by direct ablation of substrate.

3.1 Radiation pressure driven microparticle separation

Encoded microparticle which is functionalized with specific probe sequences on the surface can reduce the complexity of mixed pool of microarray oligonucleotide. Target specific primers are attached on the encoded microparticle surface and 3D dispersion of target molecules are concentrated in 2D surface by surface immobilization process. After incubation of microparticles in microarray pool for hybridization selection, they are magnetically separated. Focused pulse laser transfer radiation energy through transparent substrate that encoded particle with target sequence is separated with non-contact manner.

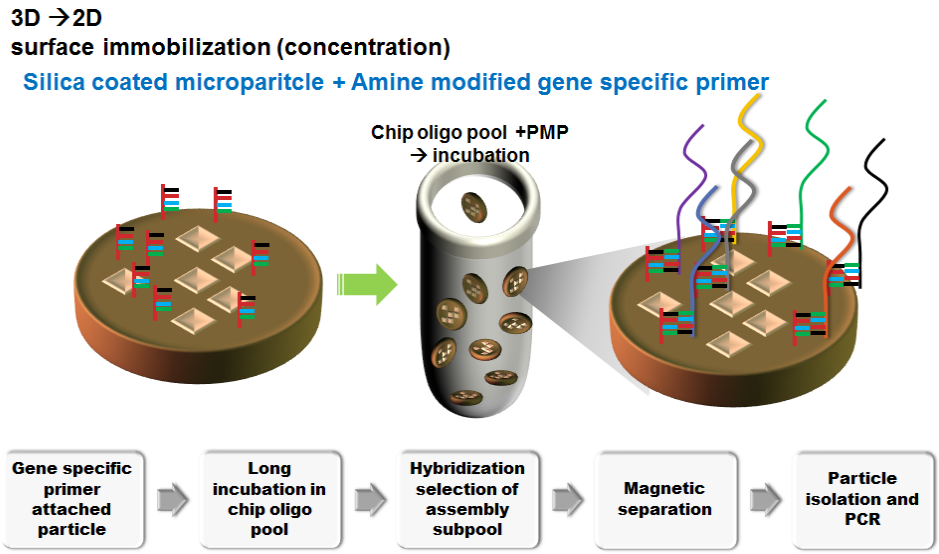


Figure 3.1 Concept diagram of probe microparticle based target molecule separation from complex pool.

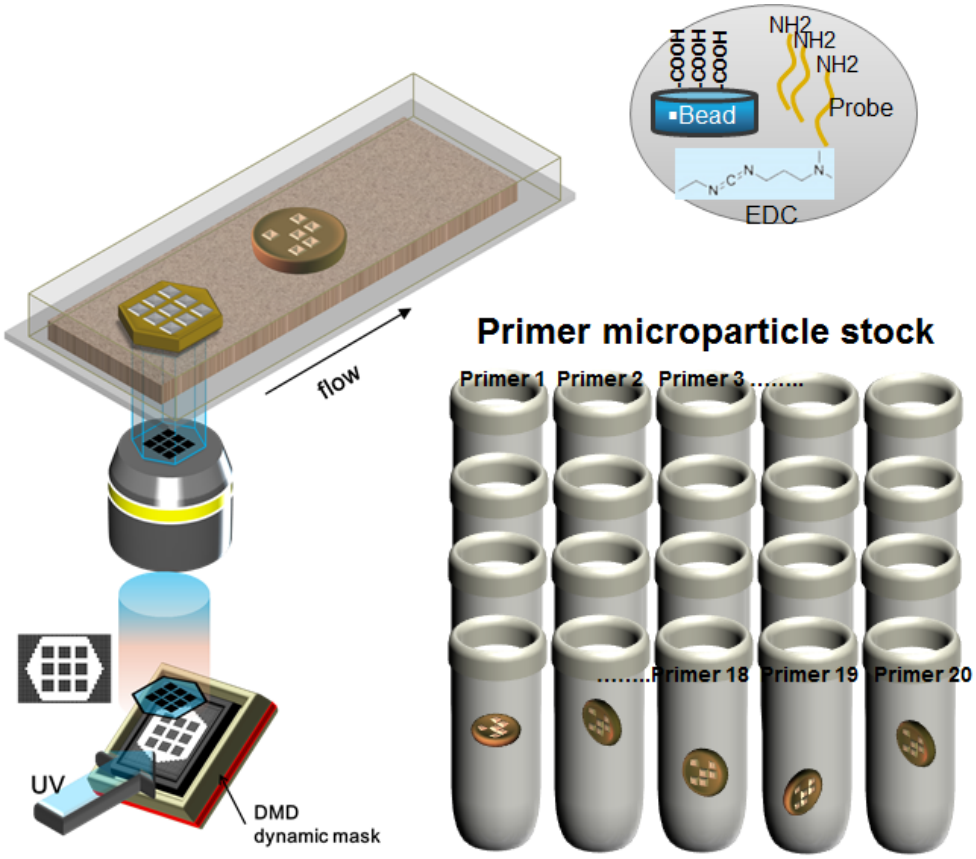


Figure 3.2 Probe encoded magnetic microparitcle fabrication using maskless lithography system. Surface of fabricated microparticles are functionalized with amine modified probe oligonucleotide after chemical surface activation.

Fabrication of surface functionalized encoded magnetic micropoarticles

Fabrication of encoded probe microparticle is composed of two sections. First, microfluidic maskless lithography system generates various shape and code microparticles [19] (Figure 3.2). Except additional precursor material for silica

surface layer, the process is almost same as those of conventional encoded magnetic microparitcle as described in chapter 2. Precursor photocurable resin (ETPTA; ethoxylated trimethylolpropane triacrylate) mixed with TMSPA (saline; for silica coating), photoinitiator and mangetic nanoparticles are immediately solidified with accordant barcode. Surface of the encoded particle are coated with silica layer by Stöber method. After series of process including amine and carboxyl group treatment, the surface is functionalized with amine modified probe oligonucleotides.

Probe sequence design

Selective separation from mixed pool of microarray DNA is achievable with orthogonal probe sequence set. Conventional orthogonal sequences are having properties of uniqueness, homogeneous annealing temperature, absence of repetitive sequence, absence of secondary structure and minimal cross hybridization [37]. Here I modify and add more filtering conditions for higher selectivity. Binding probability of each input sequence to probe sequence is derived based on matching score of Smith-Waterman algorithm and Gibbs free energy. Twenty orthogonal probe sequence data were extracted out of 50,000 randomly generated sequence pool (Figure 3.3-(a)). The probe sequence is a reverse complementary of forward primer sequence of target oligonucleotide sequence. Since the synthesis of microarray oligonucleotide starts from 3'end to 5'end, 5'end capturing of target sequence get rid

of truncated contents while synthesis process (Figure 3.3-(b)).

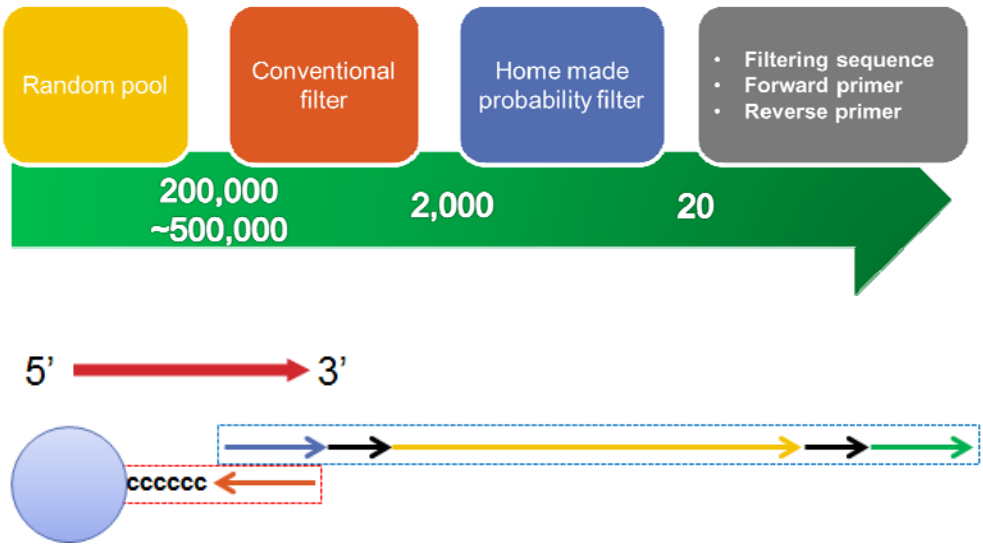


Figure 3.3 (a) Probe sequence design process. (b) Target sequence capture strategy. Amine modified probe sequence is designed for capturing target sequence in complex microarray pool. The probe sequence are a part reverse complementary of 5' end target sequence.

Radiation pressure driven microparticle separation

After incubation, probe microparticles carry target sequences from oligonucleotide pool. Following physical separation is conducted by radiation pressure driven non-contact energy transfer to avoid cross contamination as shown in Figure 3.4. Average size of microparticles are 50~100 μm that it is not trivial to physically separate them. Pulse laser energy is concentrated in the range of tens of

micrometer. Real time imaging of microparticle carrying substrate offers barcode information of target bead. At the same time, focused pulse laser is illuminated on the microparticle surface that radiation pressure separates target bead from substrate to PCR tube [25-28].

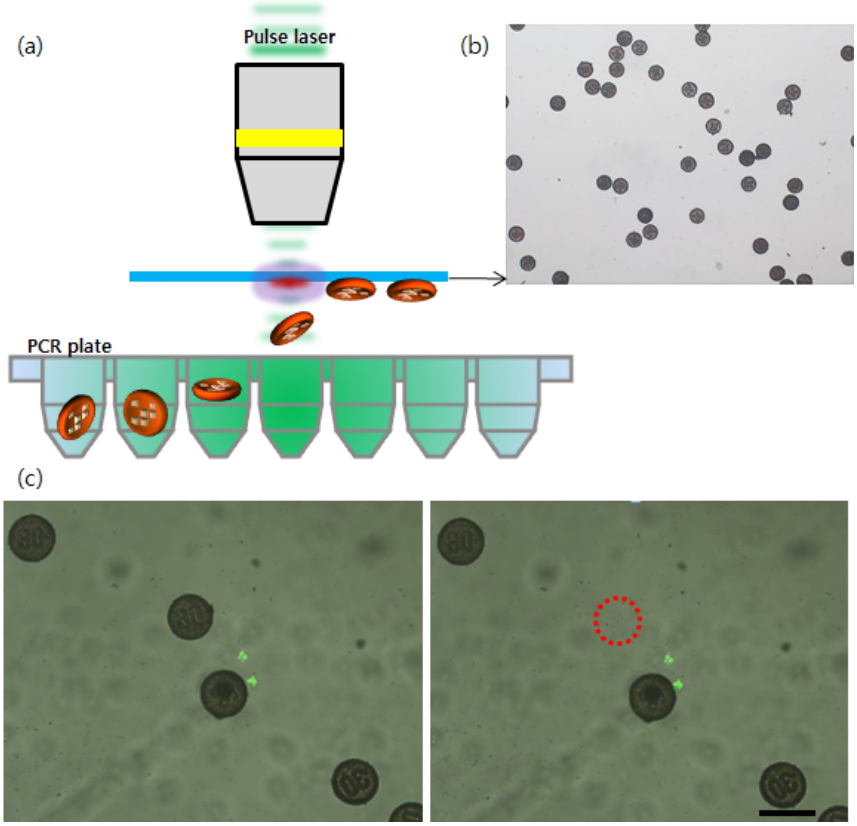


Figure 3.4 (a) Radiation pressure driven microparticle sepration. (b) Particles are encoded with numbers and dispersed on the substrate in dry condition. (c) Objective lens identify target bead location and barcode, at the same time illuminates laser pulse to target bead to detach it from the substrate.

Result

To see the correlation of capture sequence length and separation efficiency, I designed from 15mer to 35mer capture sequence and fabricated functionalized encoded particle individually. Incubated beads were isolated into PCR tube and amplified with each primer to confirm their contents. Figure 3.5 shows the gel images of PCR product of probe beads. Single bead amplification does not show significant amplification result except 35mer case. Data indicate capture efficiency of longer probe sequence is greater than shorter one. It is also can be observed in colony plate. Interestingly the vector insertion ratio is also greater when the length of probe sequence is long.

Microarray oligonucleotide pool is consists of thousands of different sequences but is categorized into small groups with same capture sequence. Those sequences are a construction unit of small gene by assembly PCR. Separated and amplified target sequences were identified by Sanger sequencing. Due to synthesis errors of microarray pool, a number of deletion and substitution bases are observed. But still probe bead successfully captured target sequences without bias (Figure 3.6)

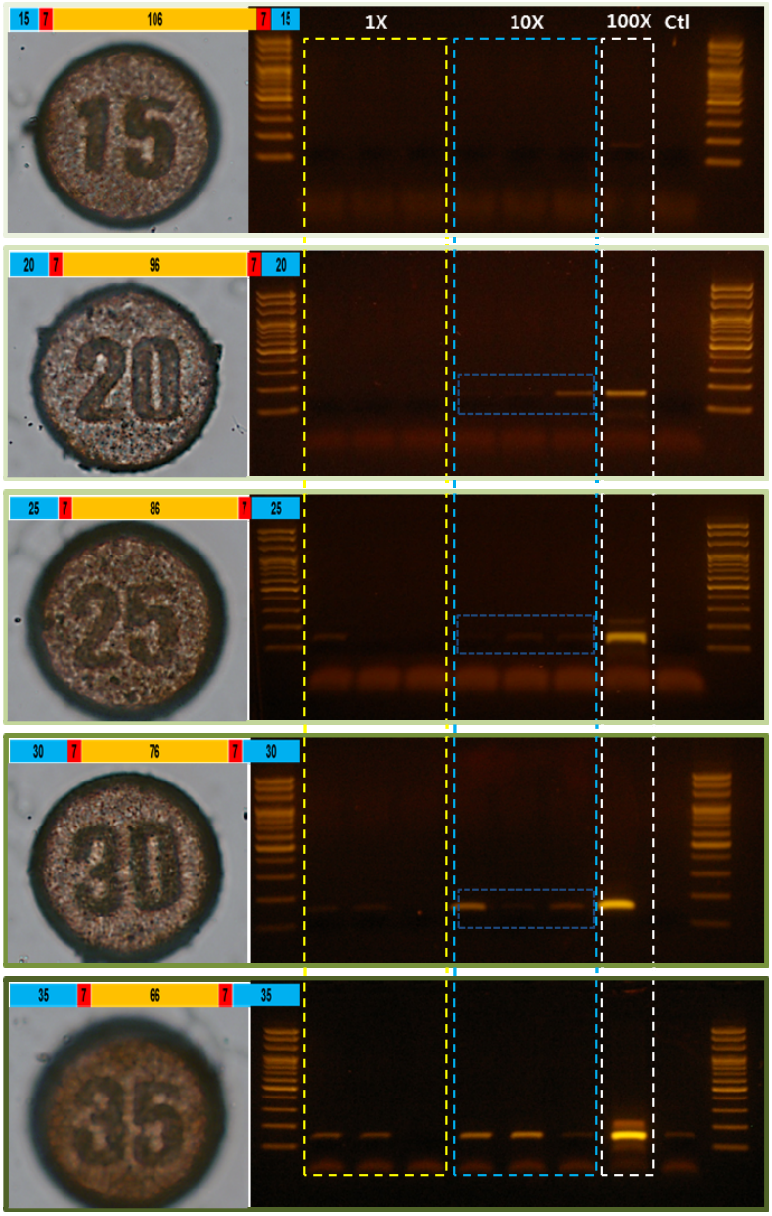
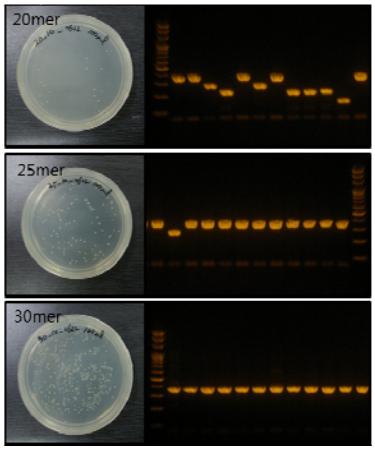


Figure 3.5 (a) Gel image of captured sequences. Longer probe sequence bead captures more target sequences out of mixed pool.

(a)



(b)

Length of flanking	Colony ID No.	Matched reference ID No.	Errors	Coverage	No. of total oligos for target gene	Error rates
20	1	5	x	3 from 5 colonies	15	83.33bp/error
20	2	8	C del, G to A, G del			
20	5	5	x			
20	7	11	T del, G to A, G to A, G del, G del			
20	12	11	A del			
25	1	12	C to T, G to A	8 from 9 colonies	17	168.75bp/error
25	3	3	T del, C del			
25	4	8	x			
25	5	13	x			
25	6	11	x			
25	7	6	C del			
25	8	16	x			
25	9	12	G to A			
25	10	7	C to T, C del			
30	1	3	x	8 from 9 colonies	19	225bp/error
30	2	3	G del			
30	3	18	x			
30	4	10	C del			
30	5	12	x			
30	6	7	G to T			
30	8	17	C to T			
30	9	15	G del			
30	10	4	T to A			

Figure 3.6 Sequence verification result. (a) TOPO cloning of captured sequence. Insertion ratio differs from probe sequence length. (b) Contents table with matched reference ID and errors.

3.2 Ablation driven separation of microarray probe

Originally, microarray is developed for high-throughput screening and diagnostic applications [29-31]. Millions of probe spots capture target sequences and generate huge amount of interaction data. In spite of its outstanding capacity, microarray has been used for identification of contents only. Here I conducted proof of concept study of direct usage of probe or hybridized molecules captured on a microarray surface. .

Separation by direct substrate ablation

Focused spot of nanosecond pulse laser generates several phenomena including heat effect zone formation, melting solid substrate into gas phase, generating debris and microcracks [32-36]. Biomolecules attached or captured on the surface of microarray can be separated with selective substrate ablation of this nanosecond pulse laser. I intentionally focus pulse laser beam at slightly beneath the target probe spot (vertical direction). Concentrated energy of pulse laser instantly change solid glass substrate into gas phase and push apart probe spot debris with target molecules as shown in figure 3.7.

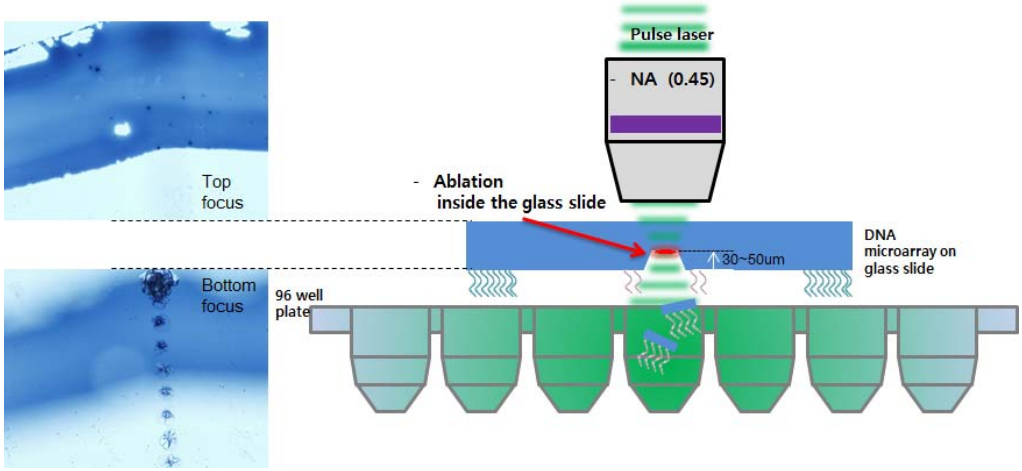


Figure 3.7 Target molecule separation via direct substrate ablation by pulse laser.

Focused pulse laser generates tens of micrometer ablation spot which is compatible with the spot size of microarray probe. I determined three reference spots by hybridization of fluorescent labeled complementary sequence (Figure 3.8).

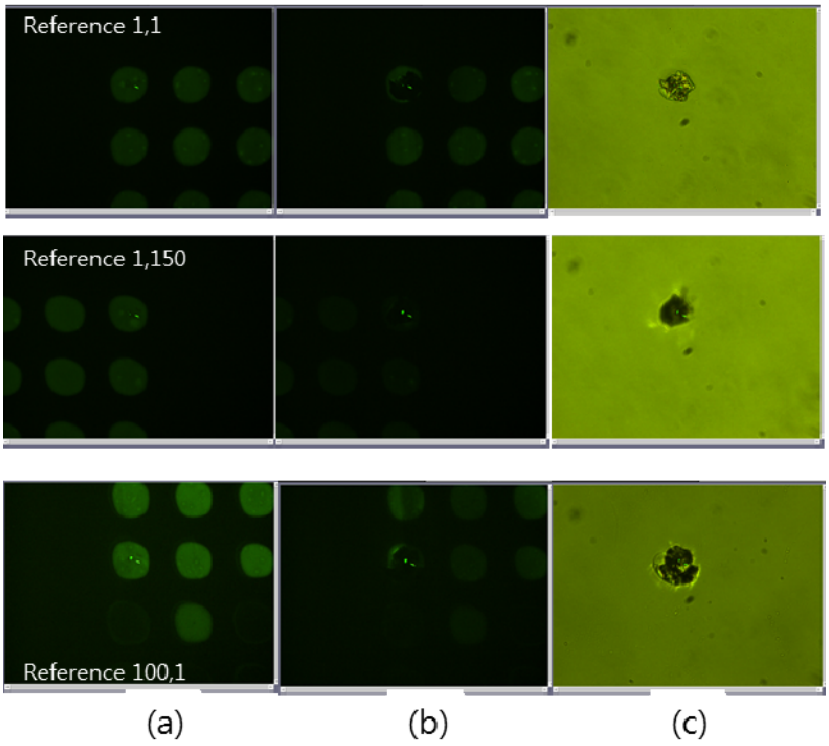


Figure 3.8 Reference spot extraction via direct ablation of microarray with focused pulse laser. (a),(b) Fluorescence label disappears after substrate ablation. (c) Bright field images of ablated spots.

According to three reference spots, motorized stage and Labview sortware

automatically locate focal spot to target probe location. Eight target spots were retrieved and successfully verified with post amplification of substrate debris (Figure 3.9).

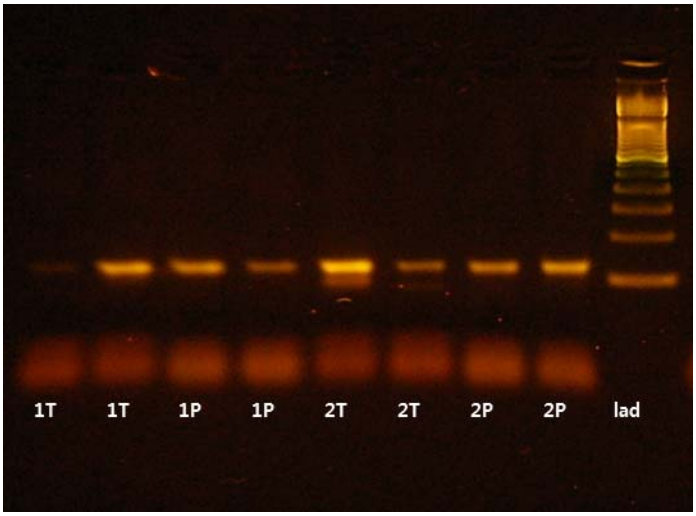


Figure 3.9 Gel image of substrate debris amplification with universal primer. Clear band indicates that substrate debris from direct ablation is carrying target molecules.

3.3 Summary

In summary, I demonstrated pulse laser driven micro structure separation technique using radiation pressure and direct substrate ablation. Non-contact nature and immediate interaction of light, target molecules are successfully separated from mixed pool without contamination. I believe this approach pave the new way to utilize rich source of microarray chip oligonucleotide.

Chapter 4

Sniper cloning method

This chapter describes the development of ‘Sniper Cloning' method using massively parallel identification followed by high-throughput deterministic separation approach to construct ultra-high quality oligonucleotide library with low cost and high-throughput manner. Process eliminates labor intensive conventional randomized clonal separation and expensive Sanger derived individual identification. I developed non-contact pulse laser retrieval system and diffusion like local mapping algorithm which enables precise and high-throughput targeting of sequence verified microbead on next generation sequencing substrate.

Writing DNA serve as numerous and significant purposes in fields of synthetic biology, functional genomics, and bioengineering. Although large amount of cheap and accurate short DNA is necessary for writing DNA as a precursor, conventional process is limited in its low-throughput and expensive nature, containing tedious randomized cloning and error elimination steps. Here I propose ‘Sniper Cloning,’ which prescreens massive amount of molecular clones in parallel, then selectively separates desired sequences for direct utilization of sequence verified contents. I developed pulse laser radiation based high-throughput and contamination free non-contact separation method that closely combines a highly parallel synthesis and sequencing technology. Algorithm allows precise targeting of desired sequence clones identified by next generation sequencing platform and a pulse laser snipes them 7000 beads per hour with 96% accuracy. Proposed technology allows us to construct high standard 5188 shRNA library at a single run. ‘Sniper Cloning’ is a significant paradigm shift from the conventional randomized process to a highly deterministic and conservative process, and may potentially serve as a universal tool in biological sciences.

4.1 Random separation and identification

Recent progress in massively parallel DNA reading technique reveals valuable genetic circuitry whose motif can be a useful resource to ‘write’ novel DNA sequences for various applications such as protein engineering, synthetic genome

studies, *de novo* biological circuit construction and DNA memory. Writing DNA generally consists of bottom up assembly of high standard chemically synthesized short precursor DNA sequences (~200bp) to avoid synthesis errors (1 in hundreds base pairs) rather than constructs at one sitting from beginning to end. Thus current writing targets having at least several mega base pairs length sequence requires over hundreds of thousands error-free precursor DNA. Although state-of-art oligonucleotide synthesis technology offers extremely rich source of DNA, downstream error elimination process; manual pick-and-place cloning separation followed by Sanger identification, limits the progress due to low throughput and high cost [35]. Since conventional molecular cloning relies on the fully randomized selection process of the clones, it severely hampers the recovery rate of error-free target sequences. Moreover, the number of picking and individual capillary sequencing should be exponentially increased when using cost effective microarray-derived mixed pool DNA, which typically shows higher error rate and PCR enriched amplification bias, further demanding the need for addressing this limitation.

Unnecessary additional work in randomized error-elimination process is avoided by using megaclones of next generation sequencing (NGS). NGS based on parallelized sequencing-by-synthesis not only offers large sequencing information but also leaves millions of sequence-verified clonal bead on substrate as soon as sequencing finishes [55]. The two lucrative contents that NGS produces can be highly

cooperative, allowing massively parallel prescreening followed by deterministic separation of target contents. Such utilization may dramatically decrease the burden of conventional randomized cloning separation and Sanger sequencing. However, numerous clones on NGS substrate are useless without precise clonal bead locator and appropriate high-throughput retrieval method. In practice, a recent physical pick-and-place approach for the retrieval of clonally amplified DNA bead on NGS substrate, mechanical instability and imperfect clone location tracking method lowers the retrieval throughput that follow-up studies have not been made until now [56].

Sniper Cloning

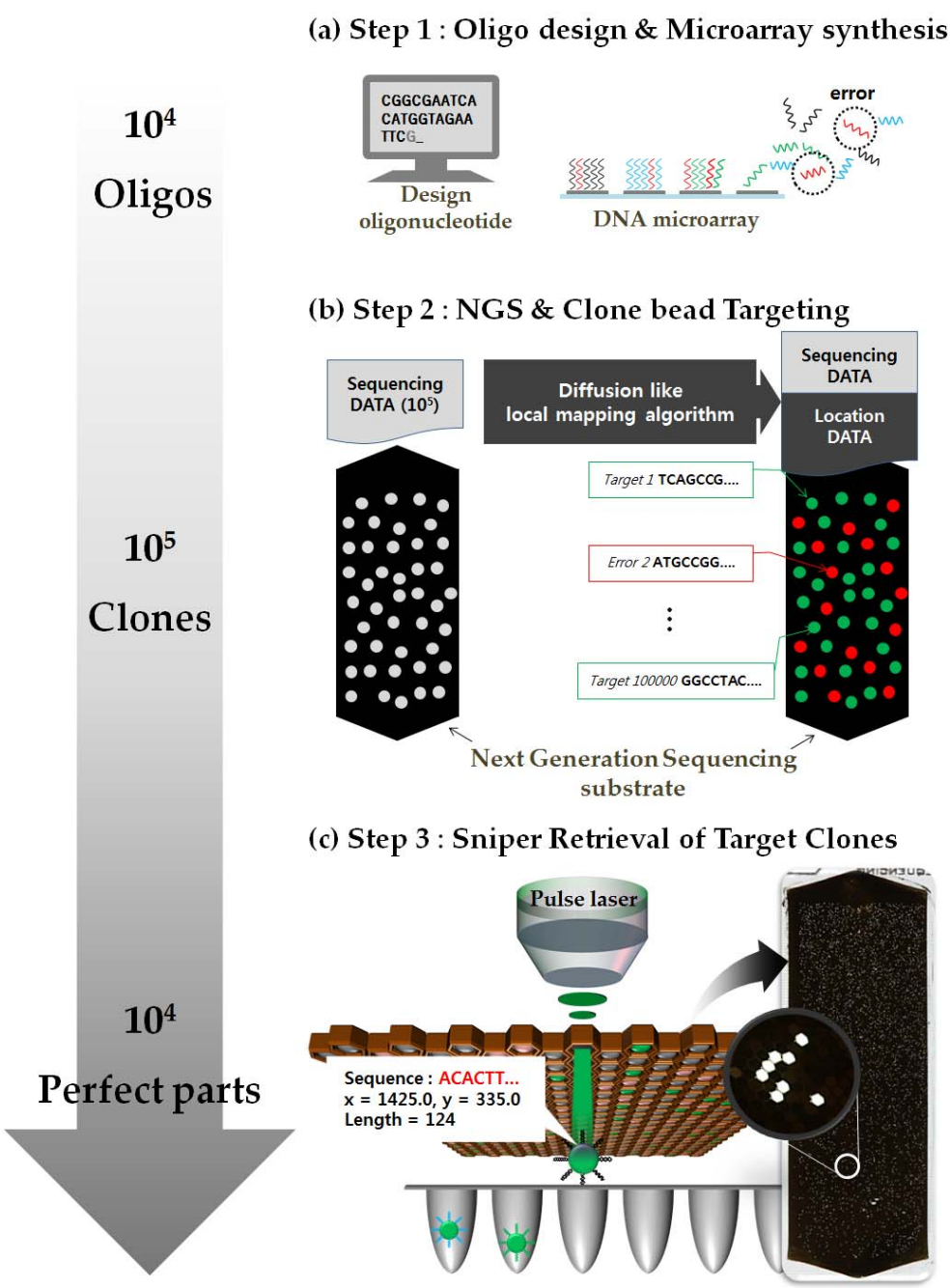


Figure 4.1 Sniper targeting of molecular clones. (a) Parallel synthesis of custom design oligonucleotides. Constructed sequences are cleaved from substrate and forms mixed pool including more than 10^4 kinds of oligonucleotides with certain portion of errors (red) at once. (b) Next generation sequencing (NGS) based massively parallel identification followed by clone bead location targeting using diffusion like local mapping algorithm. NGS isolates and amplifies single molecule from a complex pool to supply huge amount of sequencing data (10^5) with accordant pixel information. Diffusion like local mapping algorithm overcomes random and nonlinear distortion of sequencer's imaging system, and converts pixel information into real-world location of target clone bead. (c) Sniper retrieval of target clone beads. Precise location data from clone bead targeting step and non-contact pulse laser bead retrieval system allow us high-throughput (two beads per second) selective separation of 10^4 perfect parts without cross-contamination. Integrated process including parallel synthesis (DNA microarray), parallel identification (NGS) and high-throughput separation (pulse laser optical retrieval system) dramatically reduces necessary resources by eliminating randomness of conventional cloning method and provides huge amount of ultra-high quality artificial oligonucleotides within few days.

4.2 Radiation pressure driven high-throughput separation system

Sniper retrieval of target clones on NGS substrate is achievable by setting up a closely integrated technical procedure including DNA microarray, NGS, and pulse laser retrieval system. DNA microarray synthesizes more than ten thousand short (120bp) single strand oligonucleotides with certain portion of errors (Figure 4.1-(a)). The NGS platform, GS Junior from Roche 454 Life Sciences, identifies the contents of DNA complex pool from microarray through *in vitro* cloning followed by massively parallel pyrosequencing. I developed diffusion like local mapping algorithm to pin-point exact location of target clone beads on substrate, and selectively separate beads containing desired sequence verified oligonucleotides for direct utilization (Figure 4.1-(b)). I use the radiation pressure of focused pulse laser to retrieve target beads from sequencing substrate (Figure 4.1-(c)) [25-28]. Non-contact nature of light fully eliminates the potential possibility of cross-contamination, which is frequently induced by physical contact such as micro tweezers or tips. In my approach, additional washing or replacement of physical equipment are not required. Also, with the help of automated linear motorized stage, high precision of focused pulse laser provides accurate targeting of desired molecular clones with minimal variation so that the high-throughput of retrieval process is achievable (2 beads per second). This complete procedure supplies massive amount of synthetic oligonucleotide with extremely high standard.

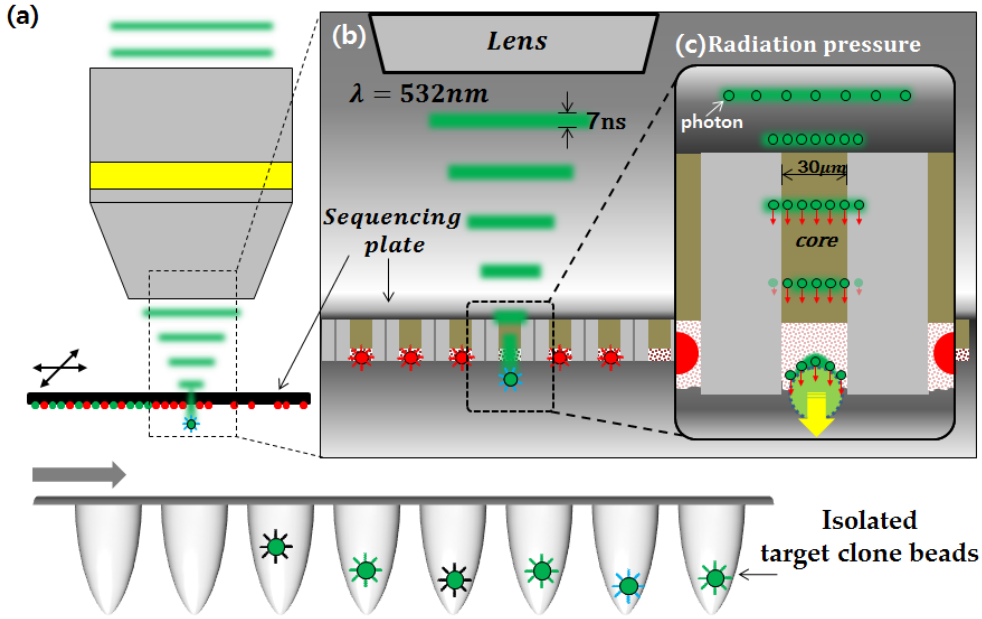


Figure 4.2 Schematic diagram of focused pulse laser radiation pressure driven non-contact target bead sniper system. (a) Motorized stage moves sequencing plate and locates target clone bead to focusing spot of pulse laser based on real-world location information from diffusion like local mapping algorithm. Target clone beads were isolated into PCR tube to directly utilize sequence verified oligonucleotides on bead surface. (b) The selectively etched fiber bundle structure of 454 NGS substrate is well suited for optical releasing. Etched core region partially isolates a single clone bead while remnant core delivers serial optical signal of pyrosequencing to CCD. Biocompatible 532nm visible light is illuminated to opposite side of bead containing side and couple them into core region to carry photon energy to target bead penetrating through sequencing plate. Nano second pulse effectively exerts radiation

force to target bead without both physical damage (longer pulse) and simple locoregional ablation (shorter pulse). (c) Since fiber only carries lights in core region and attenuates them otherwise, it minimizes the horizontal and vertical positioning error as well. It allows robust clone bead targeting without expensive optical and mechanical instruments. A single pulse having energy of around 50μJ applies 0.25 μN radiation force for target bead retrieval.

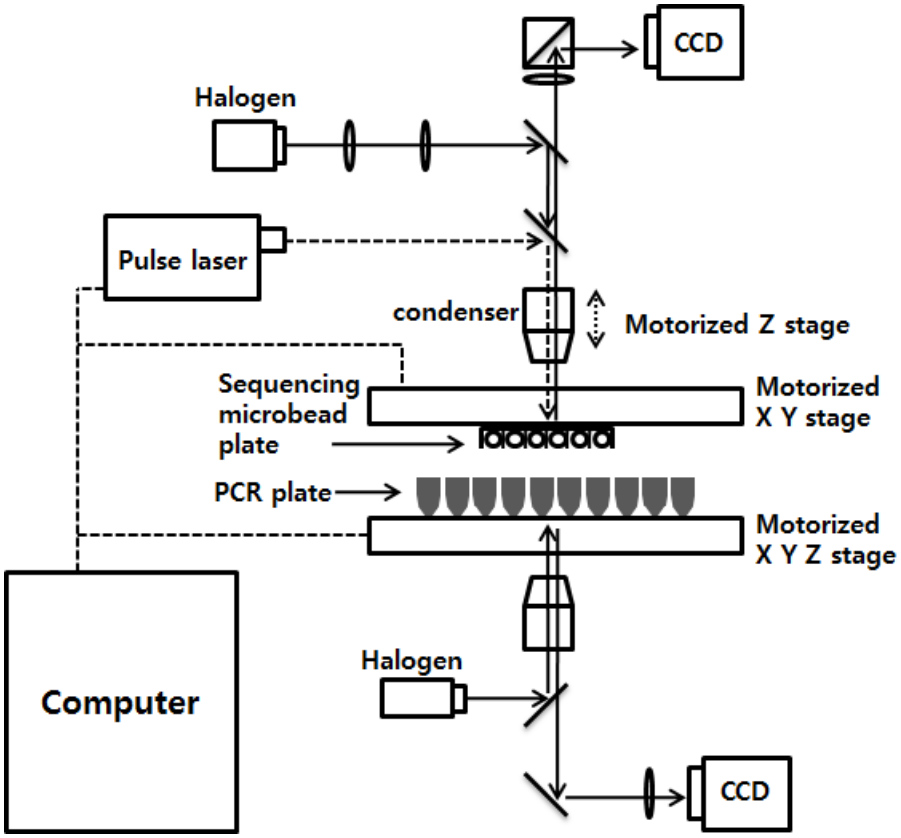


Figure 4.3 Schematic diagram of pulse laser bead retrieval system. Pulse laser, CCD camera and two motorized stages are controlled by personal computer with self-made Labview sortware. Upper part of the system, commercial inverted microscope and a motorized stage are hanging upside down that the direction of radiation force is identical to that of gravity. I constructed whole system on the anti-vibrational optical table except personal computer and pulse laser power supply.

I actively took advantages of optically favorable substrate structure of 454 Junior platform. Selective etching of fiber bundle not only serves as isolation chamber of each bead but also optical path of successive pyrosequencing information through substrate to CCD front. As depicted in figure 4.2-(b), I have inversely delivered harmless, low energy (50μJ/pulse) visible nanosecond pulse laser (532nm;7ns) to couple them into remnant core of the target well from the back side of the substrate. Remnant fiber core guides light pulse that pushes target molecular clone bead with radiation pressure. The amount of impulse can be calculated as follows

$$\text{Radiation pressure : } \frac{S(t)}{c}$$

where, $S = \frac{Energy_{total}}{Area \cdot Time}$ and velocity of light c.

$$Energy_{total} : 50\,\mu J \text{ /pulse,} \quad \text{Area : } \pi(15\mu m)^2, \quad \text{Time: 7ns}$$

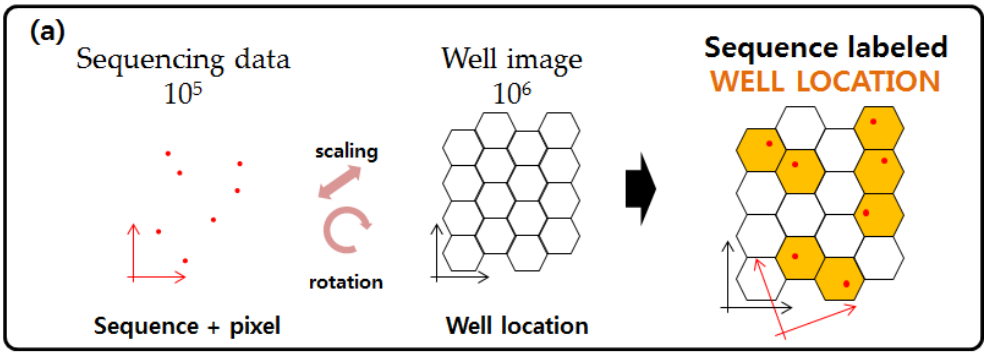
Hence within 7ns, target bead will get radiation force of $0.25\,\mu N$ that totally 1.67fNs impulse is transferred to bead from focused single pulse laser shot. Since longer pulse duration especially CW laser may boil target points and expensive shorter pulse laser (ps, fs) can cause avalanche breakdown, nanosecond pulse could be the most appropriate energy transfer method to retrieve target beads [32]. Retrieved bead molecular clones are collected into lower 96 well plate for the subsequent amplification (Figure 4.2-(a)). Since the fiber only carries lights in core region and attenuates them otherwise, it minimizes the effect of positioning and

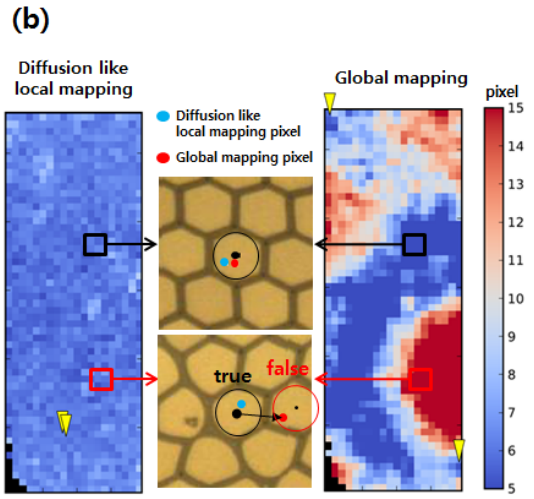
fabrication errors of both horizontal and vertical coordinates of the substrate (Figure 4.2-(c)). It makes this system robust, rid any expensive optical and mechanical instruments.

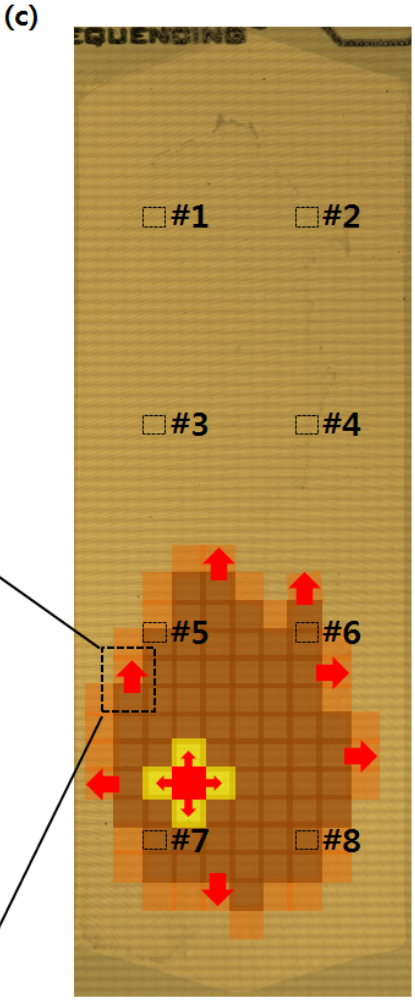
4.3 Clone tracking algorithm

The true position of target clone beads on sequencing substrate can be found by overlapping pixel map from NGS data and well center position map of stitched whole chip image (Figure 4.4-(a)). However, due to random and nonlinear distortion of sequencer's imaging system, simple linear transformation of the error-prone pixel values are not appropriate to recover precise true location of each sequence verified clonal bead throughout whole chip range [46] (Figure 4.4-(b) and 4.5). The only way to eliminate positional error induced by physical distortion is to localize the region of interest that has acceptable amount of distortion. Self-designed 'Diffusion Like Local Mapping Algorithm' divides whole chip area by 300 semi-linear subdomains with slight overlaps (Figure 4.4-(c)). After then mapping calculations between pixels and corresponding well location were diffused throughout whole chip range from one initial subdomain containing two Sanger verified reference beads by adjusting scale and rotational angle of pixel domain. Adjacent subdomains were consecutively mapped according to two new reference

points in overlapped region that were determined by mapping result of previous subdomain (Figure 4.4-(d)). Finally 10^5 sequence labeled well location data are generated out of totally 10^6 wells from stitched whole chip image.


(a)
Sequencing data
10^5
scaling
rotation
Well image
10^6
Sequence labeled
WELL LOCATION
Sequence + pixel
Well location


(b)
Diffusion like
local mapping
Global mapping
pixel
Diffusion like
local mapping pixel
Global mapping pixel
true
false
15
14
13
12
11
10
9
8
7
6
5


(c)
#1
#2
#3
#4
#5
#6
#7
#8

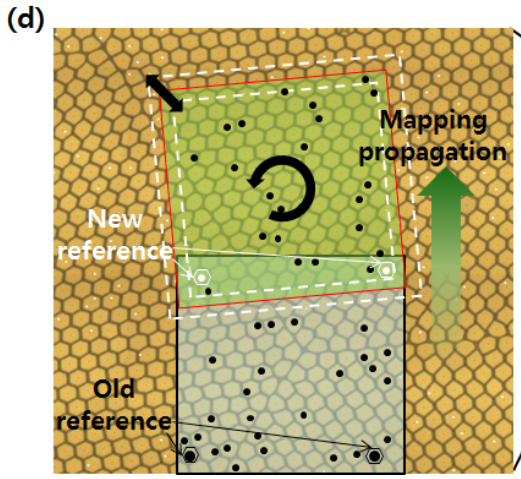

(d)
Mapping
propagation
New
reference
Old
reference

Figure 4.4 Diffusion like local mapping algorithm for detecting target bead real-world location on chip. (a) Schematic diagram of mapping algorithm. 454 Junior normally offers approximately 10^5 sequencing information with accordant pixel position of CCD. From arbitrary two reference points, corresponding sequence labeled well location can be found out by adjusting scale and rotational angle of sequencing pixel domain and overlap together. (b) However, due to random and non-linear distortion of sequencer's imaging system, one step global transformation leads to locational error. Around 20% of pixels are mapped in false position that is not distinguishable, severely dropping the reliability of whole location data. Yellow flags indicate reference points of each mapping calculation. Color bar shows pixel wise distance between mapped pixel and accordant well center. Well diameter threshold value is approximately 13.5 pixels. [59] (c) I lower the effect of imaging distortion to a negligible level by localizing region of interest. Whole chip area is divided into 300 subdomains with slight overlap. (d) One subdomain completes location mapping by supplying new two reference points to next adjacent subdomain so that local mapping propagates from initial matched subdomain through whole chip range.

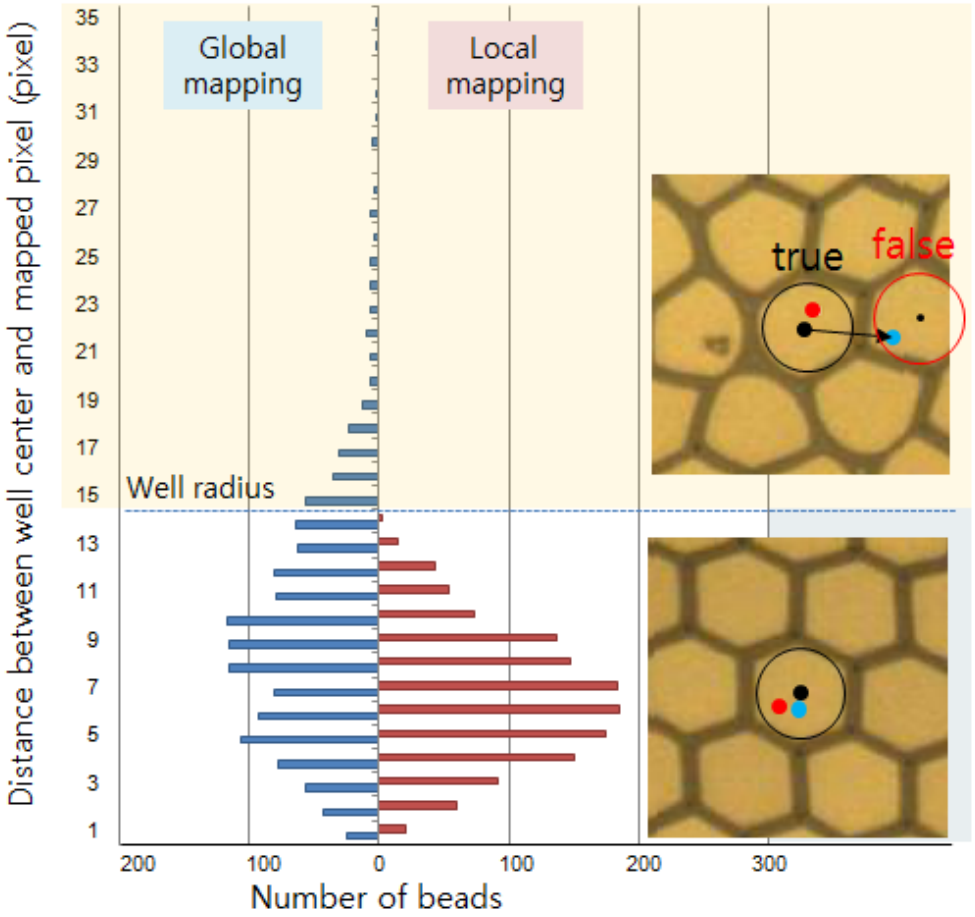


Figure 4.5 Comparison graph of conventional global mapping and local mapping algorithm. Global mapping shows 17.5% of mapped well having over threshold distance value. The false information mapped into adjacent different wells that becomes undistinguishable error.

To see the feasibility of diffusion like local mapping algorithm, I retrieved 24 target beads from 8 evenly distributed regions (Figure 4.6). Motorized stage properly moves sequencing plate to pulse laser focal point within margin of error and retrieve target beads resulting empty well at Figure 4.6-(a). Retrieved beads were amplified and verified by Sanger sequencing. Among 24 Sanger sequencing results, 20 samples have matching data from NGS results. Due to secondary structure prone characteristic of shRNA, the rest 4 samples show low quality Sanger results but still can find counterparts from NGS data using barcodes (Figure 4.6-(b)).

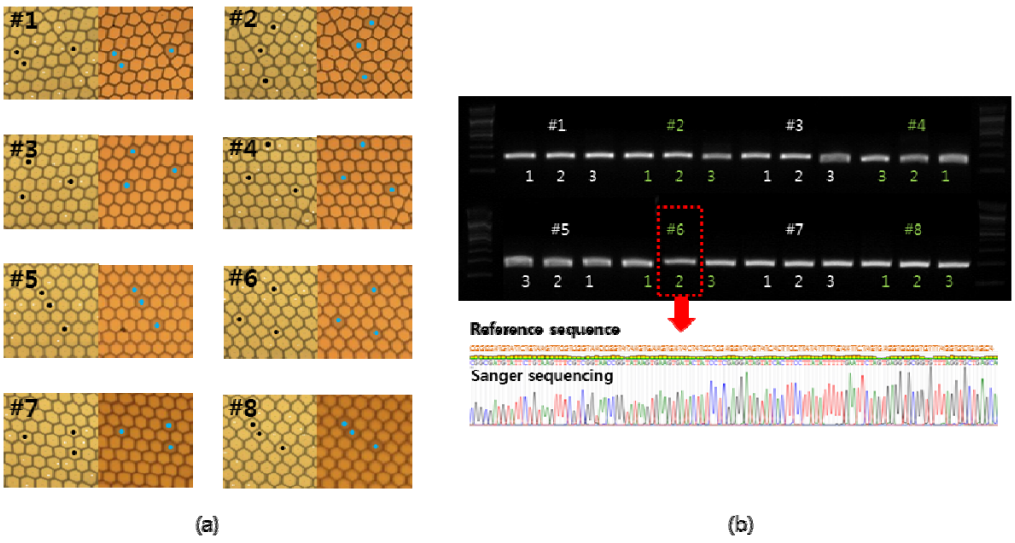


Figure 4.6 Targeting retrieval of 24 beads on eight regions. (a) 3 beads are retrieved at each region. Left side of the figure of each set indicates target location of wells and right one shows resultant empty wells. (b) All 24 beads are amplified and show clear band image. Sanger derived sequence identification confirms matched result.

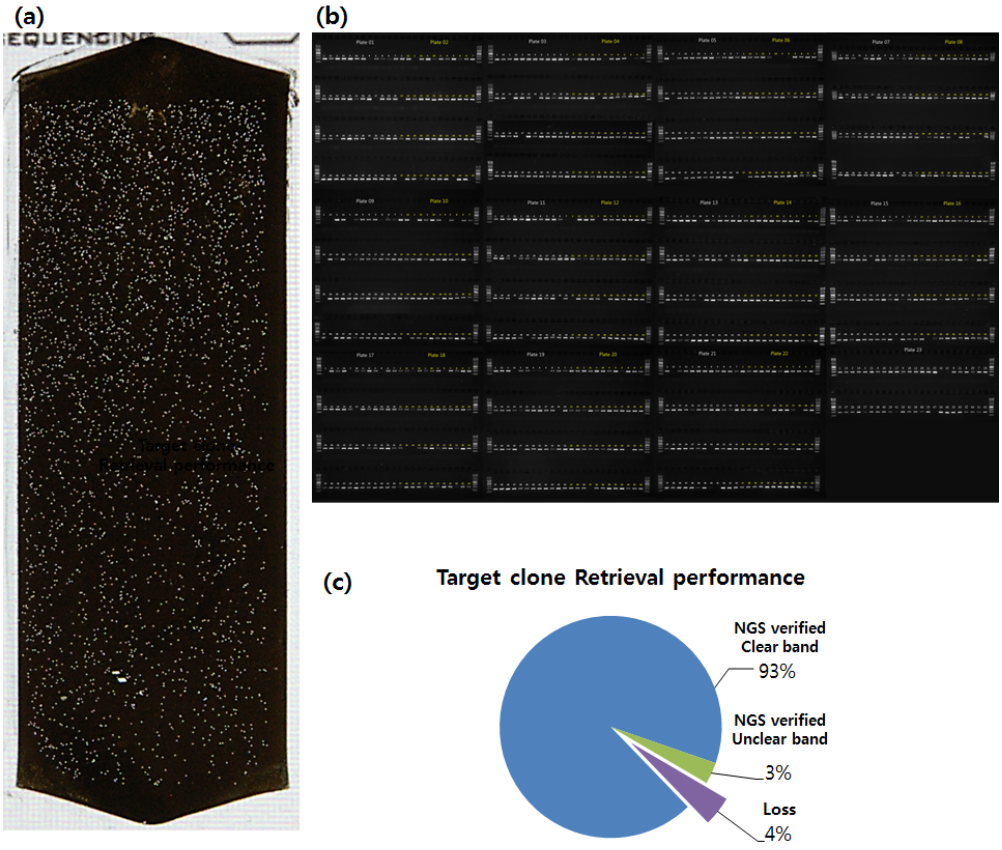


Figure 4.7 Retrieval result of 1338 shRNA sequence. (a) Transmission image of 454 junior sequencing plate after retrieval. I retrieved almost 4000 beads out of this plate. Bright spots form rectangular shape as CCD camera imaging region. (b) Gel image of all 1338 retrieved beads. (c) Amplified products are re-identified with 454 sequencing machine to verify contents. 96% of perfect match are found in verification NGS including 3% of unclear band sequence.

4.4 Result

I further verify ‘Sniper Cloning’ concept with a pool containing knockdown recombinant library by generating 1380 shRNA that targets 147 human protein (Aminoacyl-tRNA synthetases) coding genes. The sequences were synthesized by a DNA microarray [49]. After library amplification with random barcode and 454 adaptor primers, I used NGS to identify 1338 (97%) kinds of perfectly matched sequences out of 77,940 clones, and retrieved 1108 beads using pulse laser system. To acquire serviceable amount of product, additional amplification was conducted. Gel images indicate 1035(92.5%) clear band and 83 unclear or no band (Figure 4.7-(b)). Interestingly, further NGS derived sequence verification of amplified products from 1108 beads returned 1060 perfect parts (95.67%), including 58 samples out of 83 unclear bands with relatively low coverage (Figure 4.7-(c)). I believe that the 4.3% of loss originated from bead damage during sequencing run or storage, imperfect PCR condition, or retrieval process, so replicated retrieval (one sequence per at least two beads) can make loss rate to almost zero.

Target sequence structure consists of shRNA insert with primer region and restriction recognition site at both ends. To generate identical overhang sequence for cloning process, I designed four types of structure according to the insert sequence. BspE1, Age1, Xma1 and NgoMIV recognition sites are designed between front primer region and insert sequence and EcoR1 recognition site is located at the rear in

common. I attached 454 adaptors for high-througput sequencing by amplifying pool library with primers having overhang of 454 adaptor and 10mer random barcode.

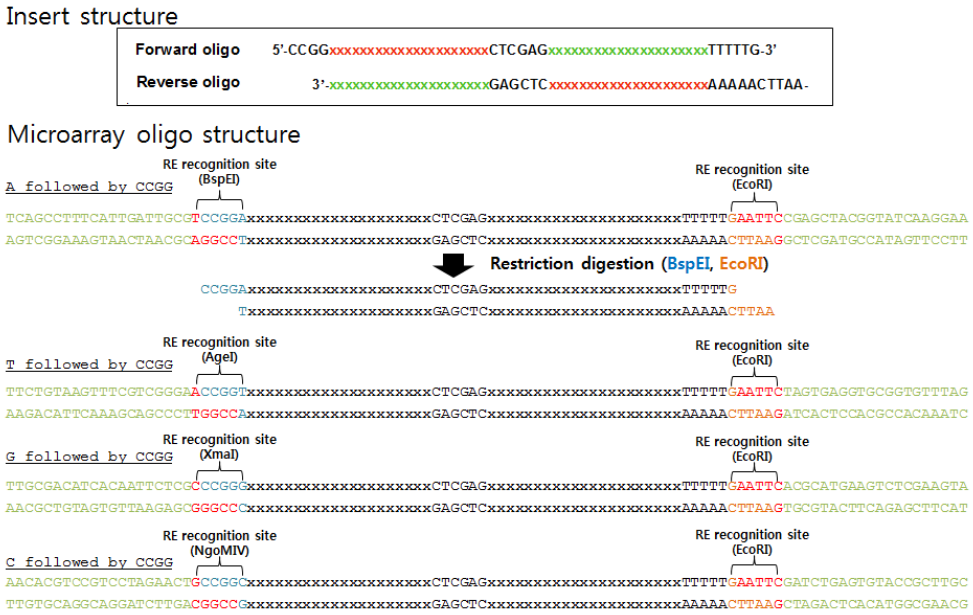


Figure 4.8 shRNA sequence structure for microarray synthesis. [60]

The effectiveness of 'identification followed by separation' approach seems to be more evident with simple mathematical comparison. Complex mixture synthesized by DNA microarray followed by library amplification generally suffers from synthesize error and amplification bias [47]. The probability (P) that a given unique DNA sequence is present in a collection of N transformant colonies is given by the expression as follows [48].

$$P = 1 - (1 - f)^N$$

where, ‘f’ is fraction of total genome.

According to this probability function, I conducted a simple calculation to find minimum number of colony picking taking account of arbitrary amplification bias and synthesis error.

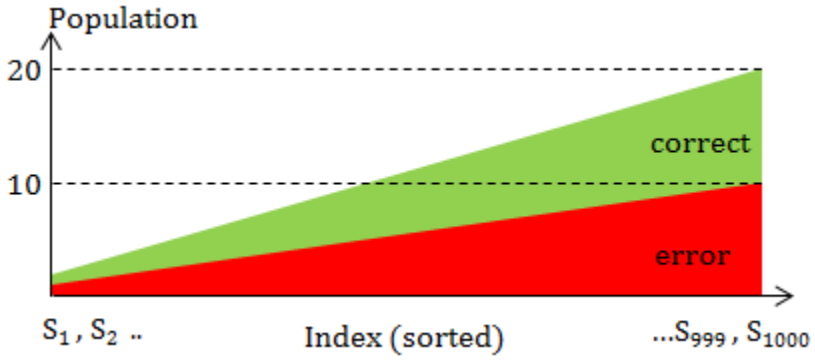


Figure 4.9 Population model with synthesis error and amplification bias.

Assume a hypothetical DNA library containing 1000 individual sequences with 50% of synthesis error and linear amplification bias (10x) as shown in Figure 4.9. After N times of colony picking, the selection probability of each sequence can be described as

$$P_{s1} = 1-(1-f_1)^N$$
$$P_{s2} = 1-(1-f_2)^N$$
$$\vdots$$
$$P_{s1000} = 1-(1-f_{1000})^N$$

where, $f_n = \dfrac{2 + \frac{18}{999} \times (n-1)}{10000}$, relative population of each content.

Thus the total probability of getting all 1000 contents is $\prod_{n=1}^{1000} P_{sn}$.

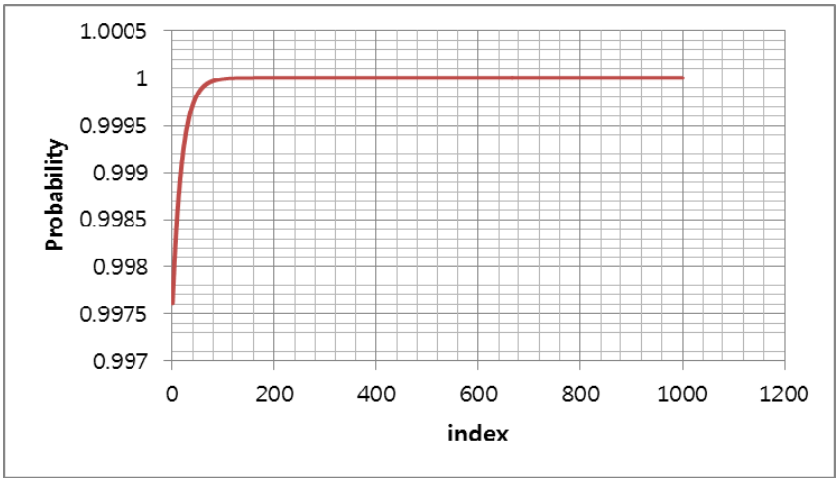


Figure 4.10 Selection probability plot of 1000 individual sequences. (N = 30197)

Figure 4.11 Total probability of getting all 1000 contents versus number of colony picking

As shown in Figure 4.11, 95.6% of 1000 sequence can be recovered when N reaches 30197 which is almost 30 times more than their original number of library contents. When one considers 50% error rate, it becomes double. To evaluate the effect of amplification bias and synthesis error, I directly compared selection probability with or without bias and error factors.

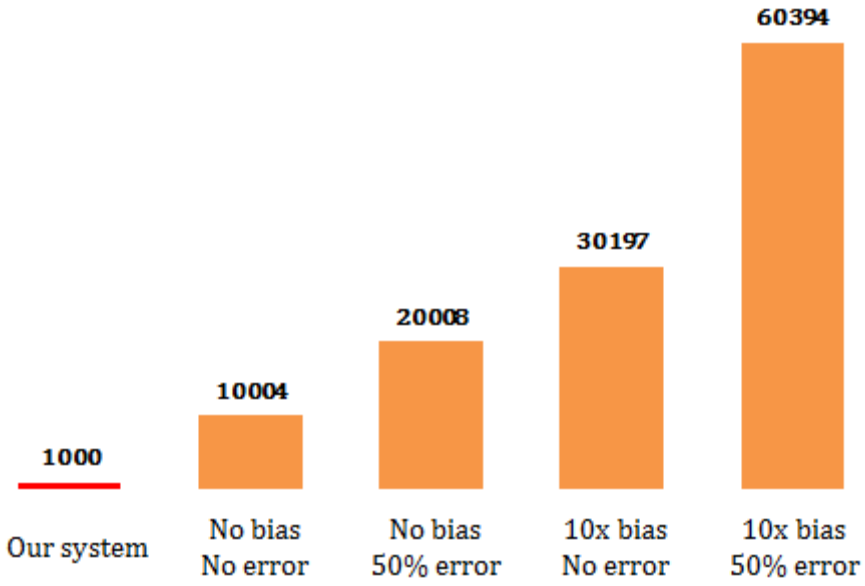


Figure 4.12 Minimum number of colony picking with or without bias and error

Among totally 330,461 shRNA clone library in RNAi consortium (TRC, Broad institute) I choose 19,061 shRNA for protein coding gene knock down (http://www.genenames.org/cgi-bin/hgnc_stats). After filtering out un appropriate sequences for pLKO1 vector cloning, I finally generate a library containing 17,959

shRNA sequences. The structure and amplification condition of second library is same as first one. 10634 sequences were randomly selected out of final 17,959 shRNA sequences.

Even if sequencing plate has only one perfect part throughout the entire population, it can be utilized without any additional work compare to large population parts. I prepared more complex DNA pool containing 10634 different sequences of human protein coding gene targeting shRNA. Although NGS result reports the relatively poor quality of microarray DNA pool and the effect of library amplification bias, I have identified 5188(48.8%) perfect parts. 99% of perfect parts have less than 10 coverage and 48% have only one copy each (Figure 4.13) I successfully separate 5188 beads from the sequencing plate within a week process including two days of microarray DNA pool synthesis, two days of library amplification and parallel identification, one day for Sanger derived reference bead determination, another two days for mapping and retrieval. Further optimization of microarray synthesis and library amplification to increase the quality of amplified pool could lead wider population coverage. Also throughput will be increased at least tenfold when it applies to 454 GS-FLX platform.

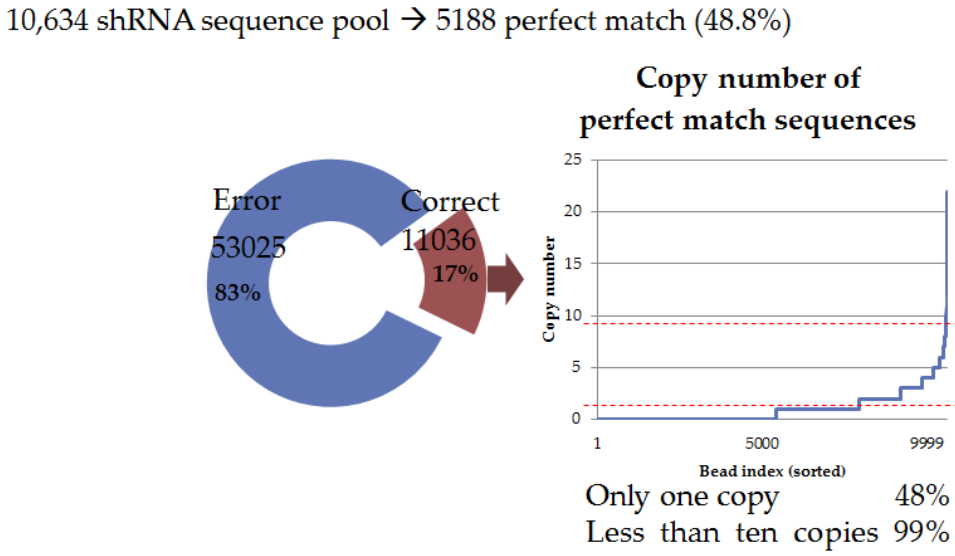


Figure 4.13 454 NGS result of 10634 pool sequence. Among 64061 number of library reads, perfect parts occupy only 17 % (11036 reads, 5188 sequences) of total read number.

To evaluate the quality of separated DNA, 454 sequencing data of 1010 retrieved bead amplicons are analyzed again. Figure 4.14 describes the portion of correct sequences in each sub-pool of different quality score reads. Red boxes show the portion of perfect reads considering only substitutional error while blue boxes take both substitutional and indel error into account. As sequencing quality score increases, the median values of blue boxes rapidly approach to that of red boxes. This means that majority of indel error reads comes from NGS sequencing error so that I can estimate the quality of retrieved DNA at least 96%.

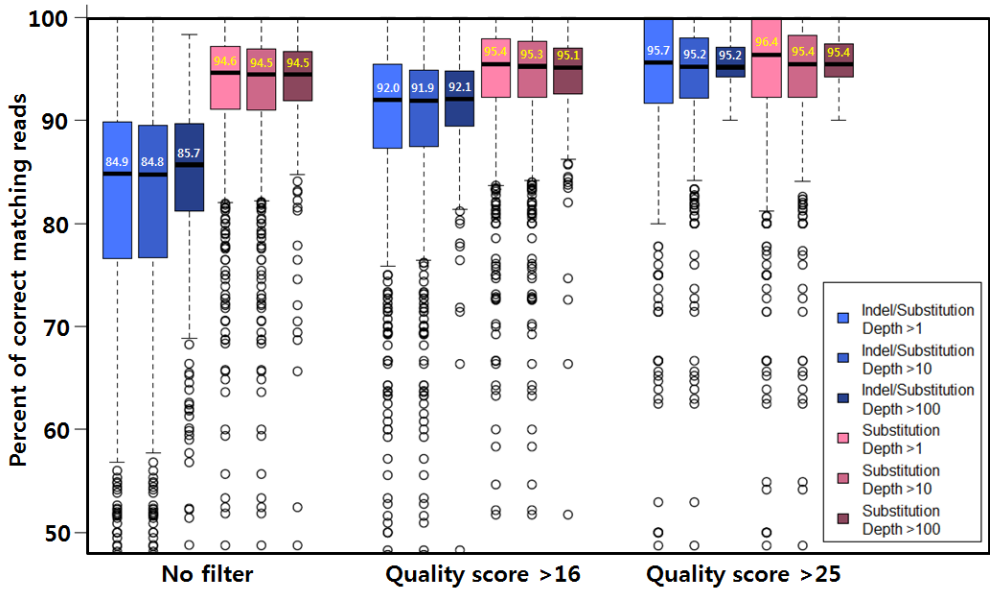


Figure 4.14 NGS based quality analysis of retrieved DNA. Box plot presents the distribution of correct reads in retrieved bead verification NGS run. 1010 kinds, totally 96484 molecules are arranged with respect to quality score and depth. Blue boxes show correct read distribution considering both indel and substitution mismatches while red boxes take only substitution into account. Median values of blue boxes in group 1 are located in mid eighty percent which is almost 10% less than those of red boxes in same group. However in group of higher quality score reads, values are gradually increased and finally reach to almost constant red box values at group 3 (95.7%). Result indicates that majority of indel mismatches are sequencing error originated from error-prone homopolymer pyrosequencing. Therefore I can say the correct portion of retrieved DNA is higher than 96%.

To the best of my knowledge, it is very difficult task to supply massive amount of synthetic oligonucleotide with extremely high standard except conventional cloning and verification. Although linear form of synthetic DNA is used in majority of cases, considerable portion of end users are still demanding circular or vector formation of DNA [51-53]. Very low error rate of products allow a direct clone-and-use strategy that eliminates following selection and identification process. I sequenced 55 shRNA inserts and found 96% (52) of the clones were perfectly correct. The rest two inserts have single base pair mismatch. Although insertion ratio varies 50~90% they still enough of a gene manipulator.

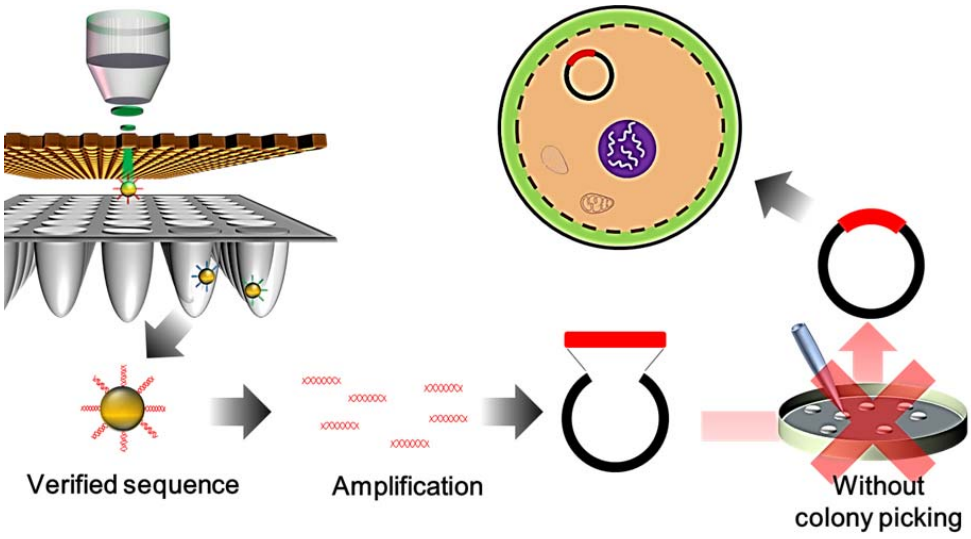


Figure 4.15 Direct clone-and-use method. Due to high quality of retrieved oligonucleotide, it can be directly clone and used without colony picking selection.

4.5 Summary

In summary, ‘Clone sniper’ approach provides massive amount of ultra-high quality synthetic oligonucleotides. Custom made pulse laser retrieval system enables non-contact contamination free high-throughput separation of perfect parts from sequencing plate with precise position data constructed by local mapping algorithm. Serial process consists of parallel synthesis, massive parallel identification and high-throughput separation dramatically reduce the necessary resources such as cost, time, labor and increase the quality of contents. I believe ‘clone sniper’ platform is able to directly utilize powerful NGS reading capacity to DNA writing ability and plays an import role on functional genome study, synthetic biology and protein engineering.

Chapter 5

Conclusion and future work

In this dissertation, I proposed and developed a series of complex biomolecule mixture separation method to fully utilize high-end parallel synthesis and high-throughput sequencing technology. A key challenge of current synthetic biology is an absence of cost effective high standard oligonucleotide precursor. Microarray DNA is a rich source of oligonucleotides that generates millions of short oligonucleotide sequence in a single run. In spite of huge advantages over conventional chemical oligonucleotide synthesis, microarray DNA, high complexity and low quality of microarray limits the progress.

In order to overcome those limitations, I developed encoded microparticle based complex pool separation technique. Maskless lithography system and

megnetochromatic ‘M-ink’ material generates magnetically manipulable color barcoded microparticles with ultra-high coding capacity. Surface of microparticles were functionalized with probe oligonucleotides that decrease complexity by capturing target oligonucleotides from complex pool.

I also developed non-contact microstructure separation technique based on focused pulse laser. Functionalized encoded microparticle and DNA microarray can lower down the complexity level to specific purpose and indirectly identify their contents. However, final physical isolation should be carried out for individual utilization. I applied radiation pressure of focused pulse laser as a microstructure non-contact manipulation tool. Focused nanosecond pulse laser spot immediately exerts expelling radiation force to microstructure attached on the transparent substrate and successfully separates them into isolation chamber. On the other hand, microarray probe spots are fixed on substrate that radiation pressure driven separation approach is not available. Hence I applied more intense focused pulse laser to slightly beneath the probe surface for direct ablation of target containing substrate. Micro-scale debris of substrate which carries probe or captured target molecules are separated and isolated for post usage. By using this pulse laser approach, high-end identification tool such as microarray becomes a rich individual source of construction.

Another important aspect of precursor oligonucleotide is the quality.

Synthesis error of microarray cannot be perfectly avoided. The most fundamental way to sort out error containing sequences is individual identification and separation. However, conventional molecular cloning separation and identification needs huge resources such as time, money and labor due to its randomness and nonparallel characteristic. To reduce necessary resources and increase throughput, I developed non-contact ‘clone sniper’ method which closely combined with parallel synthesis, parallel identification technique. Huge amount of cost effective microarray oligonucleotides are first identified by next generation sequencing platform with parallel manner. Algorithm I developed precisely tracks the location of each target sequence micro clone and following radiation force driven by focused pulse laser snipes them to separate. This ‘massive parallel identification followed by high-throughput separation’ approach eliminates the randomness of previous method and provides massive amount of ultra-high quality synthetic oligonucleotides

My final research goal is to construct *de novo* DNA writing system. Beside the rapid growth of reading technology of DNA, writing technology is limited by the absence of cost effective good precursor material and appropriate assembly method. I will extend ‘clone sniper’ method to other ultra-highthroughput next generation platforms. With novel high efficiency *in vivo* or *in vitro* assembly technique, fully automated DNA construction system has significant importance to future synthetic biology.

Bibliography

[1] Birtwell, S. & Morgan, H., “Microparticle encoding technologies for high-throughput multiplexed suspension assays,”.*Integr. Biol.*1,345-362 , 2009.

[2] Wilson, R., Cossins, A. R. & Spiller, D. G., “Encoded microcarriers for high-throughput multiplexed detection,” *Angew. Chem. Int*. Ed. 45,6104-6117, 2006.

[3] Braeckmans, K., De Smedt, S. C., Leblans, M., Pauwels, R. & Demeester, J, “Encoding microcarriers: Present and future technologies,” *Nature Rev.* Drug Dis. 1, 447-456 , 2002.

[4] Kim, H.et al, “Structural color printing using a magnetically tunable and lithographically fixable photonic crystal,” *Nature Photon.* 3, 534-540, 2009.

[5] Ge, J.et al, “Magnetochromatic microspheres: Rotating photonic crystals,” *J. Am. Chem. Soc.* 131, 15687-15694, 2009.

[6] Alivisatos, A. P., “Less is more in medicine - sophisticated forms of nanotechnology will find some of their first real-world applications in biomedical research, disease diagnosis and, possibly, therapy,” *Sci. Am*. 285, 66-73, 2001.

[7] Fulton, R. J., McDade, R. L., Smith, P. L., Kienker, L. J. & Kettman, J. R., “Advanced multiplexed analysis with the FlowMetrix(TM) system,” *Clin. Chem*. 43, 1749-1756, 1997.

[8] Han, M., Gao, X., Su, J. & Nie, S, “Quantum-dot-tagged microbeads for

multiplexed optical coding of biomolecules," *Nature Biotechnol*. 19, 631-635, 2001.

[9] Ma, Q., Wang, X., Li, Y., Shi, Y. & Su, X, "Multicolor quantum dot-encoded microspheres for the detection of biomolecules," *Talanta* 72, 1446-1452, 2007.

[10] Medintz, I. L., Uyeda, H. T., Goldman, E. R. & Mattoussi, H, "Quantum dot bioconjugates for imaging, labelling and sensing," *Nature Mater*. 4, 435-446, 2005.

[11] Braeckmans, K. et al, "Encoding microcarriers by spatial selective photobleaching," *Nature Mater.* 2, 169-173, 2003.

[12] Keating, C. D. & Natan, M. J, "Striped metal nanowires as building blocks and optical tags," *Adv. Mater.* 15, 451-454 , 2003.

[13] Nicewarner-Pena, S. R.et al, "Submicrometer metallic barcodes," *Science* 294, 137-141, 2001.

[14] Pregibon, D. C., Toner, M. & Doyle, P. S, "Multifunctional encoded particles for high-throughput biomolecule analysis," *Science* 315, 1393-1396, 2007.

[15] Dejneka, M. J. et al, "Rare earth-doped glass microbarcodes," *Proc. Natl Acad. Sci.* USA100, 389-393, 2003.

[16] Reiss, B. D.et al, "Electrochemical synthesis and optical readout of striped metal rods with submicron features," *J. Electroanal. Chem.* 522, 95-103, 2002.

[17] Ge, J., Hu, Y. & Yin, Y, "Highly tunable superparamagnetic colloidal

photonic crystals," *Angew. Chem. Int. Ed.* 46, 7428-7431, 2007.

[18] Ge, J. & Yin, Y, "Magnetically tunable colloidal photonic structures in alkanol solutions," *Adv. Mater.* 20, 3485-3491, 2008.

[19] Chung, S. E. et al, "Optofluidic maskless lithography system for real-time synthesis of photopolymerized microstructures in microfluidic channels," *Appl. Phys. Lett.* 91, 041106, 2007.

[20] Dendukuri, D., Pregibon, D. C., Collins, J., Hatton, T. A. & Doyle, P. S,"Continuous-flow lithography for high-throughput microparticle synthesis," *Nature Mater*. 5, 365-369, 2006.

[21] Yuet, K. P., Hwang, D. K., Haghgooie, R. & Doyle, P. S, "Multifunctional superparamagnetic Janus particles," *Langmuir* 26, 4281-4287, 2009.

[22] Melle, S., Fuller, G. G. & Rubio, M. A, "Structure and dynamics of magnetorheological fluids in rotating magnetic fields," *Phys. Rev. E* 61, 4111-4117, 2000.

[23] Wilhelm, C., Browaeys, J., Ponton, A. & Bacri, J. C, "Rotational magnetic particles microrheology: The Maxwellian case," *Phys. Rev. E* 67, 011504, 2003.

[24] Pregibon, D. C. & Doyle, P. S, "Optimization of encoded hydrogel particles for nucleic acid quantification," *Anal. Chem.* 81, 4873-4881, 2009.

[25] A. Ashkin and J. M. Dziedzic, "Radiation Pressure on a Free Liquid Surface," *Phys. Rev. Lett*. 30, 4,139-142 1973

[26] A. Ashkin and J. M. Dziedzic, “Optical Levitation by Radiation Pressure,” *Appl. Phys. Lett.* 19,8,283-285, 1971

[27] A. Ashkin, “Applications of Laser Radiation Pressure,” *Science* 210, 1081-1088, 1980

[28] Amol A. Ambardekar, *OPTICAL LEVITATION OF STUCK MICROPARTICLES WITH A PULSED LASER TWEEZERS,* Ph.D Dissertation, Department of Physics East Carolina University.

[29] Melissa B. Miller and Yi-Wei Tang, “Basic Concepts of Microarrays and Potential Applications in Clinical Microbiology,” *Clinical Microbio Rev,* 611–633, 2009

[30] Jörg D. Hoheisel, “Microarray technology: beyond transcript profiling and genotype analysis,” *Nature Rev*, 7, 200-210, 2006

[31] Michael J. Heller, “DNAMICROARRAY TECHNOLOGY: Devices, Systems, and Applications,” *Annu. Rev. Biomed. Eng*. 4, 129–53, 2002.

[32] B.N. Chichkov, C. Momma, S. Nolte, F. von Alvensleben, A. Tu¬nnermann, “Femtosecond, picosecond and nanosecond laser ablation of solids,” *Appl. Phys. A* 63, 109-115, 1996.

[33] M. Capitellia,b, A. Casavolaa, G. Colonnab, A. De Giacomoa, “Laser-induced plasma expansion: theoretical and experimental aspects,” *Spectrochimica Acta Part B* 59, 271–289, 2004.

[34] Alfred Vogel, and Vasan Venugopalan, “Mechanisms of Pulsed Laser Ablation of Biological Tissues,” *Chem. Rev.*, 103, 577-644, 2003.

[35] Carsten Momma, Boris N. Chichkov et al, “Short-pulse laser ablation of solid targets,” *Optics Communications,* 129, 134- I42, 1996.

[36] M. D. Perry,a) B. C. Stuart, “Ultrashort-pulse laser machining of dielectric materials,” *J.Appl.Phys,* 85, 6803-6810, 1999

[37] Qikai Xu, Michael R. Schlabach, et al, “Design of 240,000 orthogonal 25mer DNA barcode probes,” *Proc. Natl Acad. Sci.* USA 106, 2289-2294, 2009.

[38] Lodish H, Berk A, Zipursky SL, et al. *Molecular Cell Biology. 4th edition* New York: W. H. Freeman, 2000. Section 7.1.

[39] Borovkov, A.Y., Loskutov, A.V., et al, “High-quality gene assembly directly from unpurified mixtures of microarray-synthesized oligonucleotides,”. *Nucl. Acids Res.*, 38, e180, 2010.

[40] Kim, H., Jeong, J. and Bang, D. “Hierarchical gene synthesis using DNA microchip oligonucleotides,” *J. Biotechnol.*, 151, 319-324. 2011

[41] Bang, D. and Church, G.M. “Gene synthesis by circular assembly amplification,” *Nat. Methods*, 5, 37-39, 2008.

[42] Kosuri, S., Eroshenko, N., et al “Scalable gene synthesis by selective amplification of DNA pools from high-fidelity microchips,” *Nat. Biotechnol*., 28, 1295-1299, 2010.

[43] Saaem, I., Ma, S., Quan, J. and Tian, J. “Error correction of microchip synthesized genes using Surveyor nuclease,” *Nucleic Acids Res*, 40, e23 .

[44] Matzas, M., Stahler, P.F, et al, “High-fidelity gene synthesis by retrieval of sequence-verified DNA identified using high-throughput pyrosequencing,” *Nat. Biotechnol*., 28, 1291-1294. 2010.

[45] Quan, J., Saaem, I., et al, “Parallel on-chip gene synthesis and application to optimization of protein expression,” *Nat. Biotechnol*., 29, 449-452, 2011.

[46] A.R. Faruqi, R. Henderson, S. Subramaniam1, “Cooled CCD detector with tapered fibre optics for recording electron diffraction patterns,” *Ultramicroscopy* 75, 235-250, 1999.

[47] Daniel Aird, Michael G Ross, at el, “Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries,” *Genome Biology,* 12:R18, 2011.

[48] Louise Clarke and John Carbon, “A Colony Bank Containing Synthetic Col El Hybrid Plasmids Representative of the Entire E. coli Genome,” *Cell*, 9, 91-99, 1976.

[49] Michael C Bassik1, Robert Jan Lebbink, et al, “Rapid creation and quantitative monitoring of high coverage shRNA libraries,” *Nat. Methods*, 6, 443-447, 2009.

[50] Michele A Cleary\, Kristopher Kilian, et al, , “Production of complex nucleic acid libraries using highly parallel in situ oligonucleotide synthesis,” *Nat.*

Methods, 1, 241-248, 2004

[51] Jason Moffat, Dorre A. Grueneberg, et al, "A Lentiviral RNAi Library for Human and Mouse Genes Applied to an Arrayed Viral High-Content Screen," *Cell* 124, 1283–1298, 2006.

[52] Patrick J. Paddison1, Jose M. Silva1, et al, "A resource for large-scale RNA-interference-based screens in mammals," *Nature,* 428, 427-431, 2004.

[53] David E Root, Nir Hacohen, et al, "Genome-scale loss-of-function screening with a lentiviral RNAi library," *Nat. Methods,* 3, 715-719, 2006

[54] Mostafa Ronaghi, "Pyrosequencing Sheds Light on DNA Sequencing," *Genome Res*, 11, 3-11, 2001

[55] Michael L. Metzker, "Sequencing technologies -the next generation," *Nature Rev Genetics*, 11, 31-46, 2010

[56] Mark Matzas, Peer F Stähler, et al, "High-fidelity gene synthesis by retrieval of sequence-verified DNA identified using high-throughput pyrosequencing," *Nat. Biotechnol,* 28, 1291-1295, 2010.

[57] Jianping Ge, Howon Lee, at el, "Magnetochromatic Microspheres: Rotating Photonic Crystals," *J.Am.Chem,Soc*, 43, 15687–15694, 2009.

[58] Howon Lee, Junhoi Kim, at el, "Colour-barcoded magnetic microparticles for multiplexed bioassays", *Nat Mats* 9, 745-749, 2010.

[59] S. Kim, System and Algorithm for High-Throughput Retrieval of DNA Clones

after Next-Generation Sequencing, M.S. Thesis, Interdisciplinary Program for Bioengineering, Seoul National University, August 2013.

[60] H. Kim, High-throughput retrieval of molecular clones for polynucleotide synthesis, Ph. D. Thesis, Department of Electrical Engineering and Computer Science, Seoul National University, December 2012.

국문 초록

최근 합성생물학 분야의 가장 큰 화두는 ‘대량’이다. DNA 를 대량으로 빠른 시간에 읽어내는 차세대 염기서열 분석 기술의 급격한 성장은 다양한 종의 주요 유전회로 정보를 밝혔으며 이는 새로운 기능의 염기 서열을 ‘쓰기’ 위한 유용한 단초를 제공해 주었다. 염기서열 ‘쓰기’ 기술은 단백질 공학, 합성 유전체 연구 및 새로운 생물학적 회로의 구성뿐만 아니라 DNA 메모리와 같은 다양한 차세대 분야에서 응용이 가능하여 그 중요성이 더해지고 있다. 일반적으로 염기서열 ‘쓰기’ 기술은 화학 합성 시 발생하는 오류로 인하여 충분한 길이의 염기 서열을 한번에 합성하는 것은 확률적으로 불가능하며 서열이 확인된 짧은 염기서열을 상향식으로 조립해 나가는 것이 합리적인 접근 방법이다. 따라서 저렴한 양질의 염기서열 재료를 대량으로 준비하는 것이 반드시 필요하다.

마이크로 어레이 DNA 와 같은 최신의 합성 기술은 한 번에 수십만 종 이상의 짧은 염기서열을 합성할 수 있으나 오류가 섞여 있으며 원천적으로 혼합되어 있어 목적 염기서열로 조립하는데 사용하기 위해서는 적절한 대량의 식별 및 분리기법이 수반되어야 한다. 본 논문에서는 미세 입자 기술, 펄스 레이저 기술 및 위치추적 알고리즘 기술을 기반으로 차세대 염기서열 분석 장비를 통해 병렬적으로 분석된 수백만 개의 단분자 클론을 고속으로 분리하여 직접

사용하게 함으로써 마이크로어레이 칩 DNA 의 효용성을 극대화 하는 기술을 소개한다

특정 탐침으로 표면이 기능화된 미세입자는 3 차원 액상에 존재하는 혼합된 염기서열 집단으로부터 목적 염기서열만을 탐침과의 혼성화를 통해 분리한다. 코드화된 미세 입자는 혼성화 이후 펄스 레이저의 광압을 이용하여 물리적으로 분리하여 사용하도록 한다. 그러나 미세입자를 이용한 분리 기법은 동시에 여러 종류의 목적 염기서열을 3 차원 반응상에서 분리할 수 있다는 장점에도 불구하고 목적 서열에 따라 개별적으로 미세입자를 준비해야 한다는 단점이 있다. 이러한 단점을 극복하기 위하여 DNA 마이크로 어레이의 탐침 영역을 펄스레이저 식각 기술을 응용하여 직접 분리하는 기술을 제시한다. DNA 마이크로 어레이는 한 번에 수십만 종 이상의 탐침영역을 표면에 기능화시킬 수 있으며 혼성화 반응을 통해 병렬적으로 혼합된 염기서열 집단을 분리할 수 있어 효율적이다. 다만 기판으로부터 물리적으로 분리하는 방법의 부재로 인하여 관찰의 용도로만 사용하여 왔다. 본 논문에서 제안한 펄스레이저 식각 기술은 탐침 및 탐침과 혼성화 분리된 목적 염기서열을 포함하는 기판영역 전체를 물리적으로 분리할 수 있도록 하여 '쓰기'의 재료로 이용 가능하게 한다.

한편 탐침 서열로 혼성화 하여 분리된 목적 염기서열들은 합성할 때 발생한 오류를 가지는 서열을 그대로 지니고 있다. 정확한 목적 염기서열만을 분리하기

위하여 전통적으로 클로닝 기법이 사용되는데 이는 근본적으로 무작위 추출 및 뒤따르는 개별 염기서열 분석을 기반으로 하고 있어 효율적이지 못하다. 본 논문에서 제안하는 'Sniper cloning' 기법은 펄스레이저 광압 분리 기술과 정교한 위치추적 알고리즘을 바탕으로 차세대 염기서열 분석장치를 이용하여 사전 검열된 막대한 양의 클론 미세입자들로부터 목적 염기서열을 고속 비 접촉 방식기술을 이용하여 선택적으로 분리하여 무작위 추출로 인해 발생하는 불필요한 노동력과 비용을 제거하였으며 매우 높은 순도의 전구 염기서열을 대량으로 생산한다.

본 논문에서는 염기서열 '쓰기' 기술이 가지는 잠재적 발전 가능성과 이를 뒷받침 하기 위해 개발된 다양한 기반기술에 대하여 소개하고자 한다.

주요어 : 마이크로 어레이 DNA, 미세입자, 펄스 레이저, 차세대 염기서열 분석, 클로닝

학번 : 2008-30130