Ph.D. Dissertation

Automated cell sorting for spatial omics

공간 오믹스를 위한 자동 세포 분리 분석

February 2023

DEPARTMENT OF ELECTRICAL AND

COMPUTER ENGINEERING

SEOUL NATIONAL UNIVERSITY

SUMIN LEE

Automated cell sorting for spatial omics

지도 교수 권 성 훈

이 논문을 공학박사 학위논문으로 제출함
2022 년 12 월

서울대학교 공과대학원
전기정보공학부
이 수 민

이수민의 공학박사 학위논문을 인준함
2022 년 12 월

위 원 장 정 준 호 (인)

부위원장 권 성 훈 (인)

위 원 김 진 홍 (인)

위 원 최 홍 윤 (인)

위 원 정 연 철 (인)

Abstract

Spatial omics profiling technologies have been recognized recently for its ability to decipher the genetic molecules that are structurally relevant in pathology. Especially, in tumor biology, tumor is not the group of malignant tumor cells, but rather group of various cells such as tumor cells, immune cells, fibroblasts, etc. gathers together, constructing the tumor microenvironments. Technologies to analyze such microstructures have evolved from bulk sequencing, single cell sequencing to spatial omics profiling technologies. Spatial omics profiling technologies have highly influenced in decoding cancerous mechanisms by questioning the tumor heterogeneity, tumor microenvironment and spatial biomarkers.

Most of the spatial omics technologies focus on mapping the spatial omics landscape in a large scale. They rather introduces the spatially-barcoded capture probes or fluorescence labeled target probes to spatially locate the genetic molecules. The information depth and the scalability of the techniques varies according to the

purpose of the spatial assay techniques. Such technologies are capable of discovering the spatial heterogeneity and the spatial landscape of the consisting cell types due to relatively low depth of the omics information. To effectively address the target molecules for therapeutics or diagnostics, higher depth of the omics information are required. To meet the needs, region of interest (ROI)-based spatial technologies isolated the target regions and applies chemistries for higher coverage omics data.

Conventional cell sorters utilizes microfluidic channels to sort cells of interest which requires cell dissociation in a solution phase. For instance, Fluorescence activated cell sorter (FACS) or Magnetic-activated cell sorting (MACS) uses fluorescence or magnetic particles, respectively, to designate the cells of interest in dissociated cell solutions. Spatially isolating techniques such as laser capture microdissection (LCM) is able to sort out the ROIs while preserving the spatial context, but it approximately takes an hour for isolating the targets. Also, it uses rather UV laser to dissect out cells or IR-activated melting of polymers to stick out cells which might cause damage to cells.

Here, I developed the automated spatially resolved laser activated cell sorter that isolates the cells in target per second while preserving the spatial context of the cells. Specific region of indium tin oxide (ITO) coated slide glass evaporates when illuminated by IR laser pulse, plunging the cells into the desired reservoir. The applicability of the suggested cell sorter are demonstrated in omics profiling chemistries such as DNA sequencing, RNA sequencing, mass spectrometry, etc.

Keywords: Spatial omics, Spatially resolved laser-activated cell sorting (SLACS), Laser capture microdissection, Image processing

Student Number: 2018-24629

Table of Contents

List of Tables

List of Figures

Chapter 1.

Introduction

In this chapter, recent developments of the spatial omics profiling technologies and its application to cancer biology is presented. Novel discoveries in tumor biology have been performed by tools to analyze cellular biology such as bulk sequencing and single-cell sequencing. Tumor forms the tumor architectures consisting of group of various types of cells, affecting the underlying cancerous mechanisms. Therefore, it is important to decipher the cancer while preserving the spatial context. The spatial omics profiling technologies can be mainly categorized into spatial landscape mapping technologies and ROI profiling technologies. The purpose of each technologies differs according to the cell throughput and information depth of the omics profiles. Novel techniques to map tissues in spatial context or ROI-based omics profiling have greatly discovered the oncological problems in terms of tumor heterogeneity, tumor microenvironment, or geographically localized biomarkers.

1.1. The Need for Spatial Profiling in Cancer Biology

1.1.1. Historical Review of Technologies to Address Cancer Research

Cancer biology is similar to astronomy studying an unknown universe in terms that despite the knowledge that we have stacked, there are astronomical amount of researches remaining. Looking back in the history of astronomy, the field has grown along with the advances in tools to interpret space. Ancient astronomers mapped the constellations by staring at the dark sky with bare eyes. The desire to better stare at the stars led to the development of the telescopes, from Galileo Galilei, Satellites, and Hubble telescope all the way to today's James Webb. And based on these technological developments, the findings matured from simply identifying the moons of the planets in the past, discovering new extrasolar planets and measuring space expansion rate in the 2000s, to gathering more details on black hole4 and analyzing the space atmosphere in recent years. Each telescope has a different purpose leading to different scientific findings, but they share the ultimate goal of understanding the universe: where we are, and from when we were.

Similarly, when considering the cancer history, the first definition of cancer as a disease have been proceeded by ancient biologists as early as 2500 B.C. and they even tried to eliminate it via surgery. However, due to the lack of tools to decipher the cellular analysis, the cancer was described as a disease which is caused by an excess of the black bile [1]. It was 17th century when the cells were first identified under the microscope by Robert Hooke [2]. Johannes Müller then characterized the

abnormal phenomenon in cancer cells by observing cells under the microscope [3]. He was able to define the carcinoma subtypes suggesting the heterogeneity of cancer[4]. In the mid-19th century, Rudolph Virchow have discovered the cancer cells were from normal cells, and published the Microscopic Atlas of Pathology, suggesting the various cancer types [4,5].

Microscope itself has greatly contributed in cancer research, but thrived in the aid of the various staining methods. Joseph von Gerlach first came up with the concept of staining tissues with carmine [6], which eventually contributed in the development of hematoxylin and eosin (H&E) staining [7]. H&E staining method is still used these days to simply identify the cellular organization [8]. Additional methods such as immunostaining methods or fluorescent in situ hybridization (FISH) have been developed to label the specific cells of interest. It was in 1941 that immunostaining methods was first developed using the antigen-antibody interactions [9,10] and in 1969, the in situ hybridization methods was first suggested to target genetic molecules under microscope [11]–[13]. Both methods have highly affected in stacking the knowledge of the cellular mechanisms of cancer and heterogenetic landscapes [14].

Beside from visual examination of the cellular phenotypes, other methods to profile the genetic molecules have been developed such as next generation sequencing methods. After the Human genome project which has outputted the full genome sequence of the human, it has been suggested that the cancer cells might have been derived from the genetic mutation of normal cells [15],[16]. In 2005, the

cancer genome atlas project was launched to obtain the cancer-related genetic sequences and was completed in 2018 [17]. It has been highly recognized that cellular heterogeneity exists in cancer cells, and the single cell sequencing technologies have also been developed to answer the composition of tumor cells, intratumoral heterogeneity and tumor progression [18]–[20].

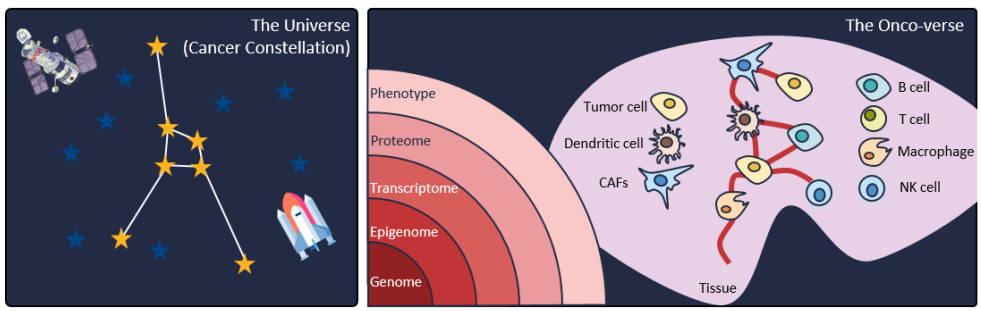


Figure 1.1 Cancer biology is similar to the universe

Through the advances in technologies to decipher the tumor biology, it has been discovered that the tumor constructs the tumor architecture resulting in the underlying cancerous mechanisms [21]. Growing number of research insists the importance in analyzing the tumor architectures in spatial context, questioning the mechanisms under tumor initiation, progression, evolution, metastasis and following diagnostics and therapeutics strategies. It is reported that various cell types affect tumor cells, constructing the tumor microenvironment consisted of B cells, T cells, macrophage, dendritic cells, fibroblasts, etc. Clinical outcome is greatly relevant to the spatial composition of the tumor microenvironment (TME) and cancer subtypes [22]–[27]. The need for the spatial analysis of tumor architecture have been

emphasized to analyze the tumor heterogeneity, TME, and spatial biomarkers[27]-[33].

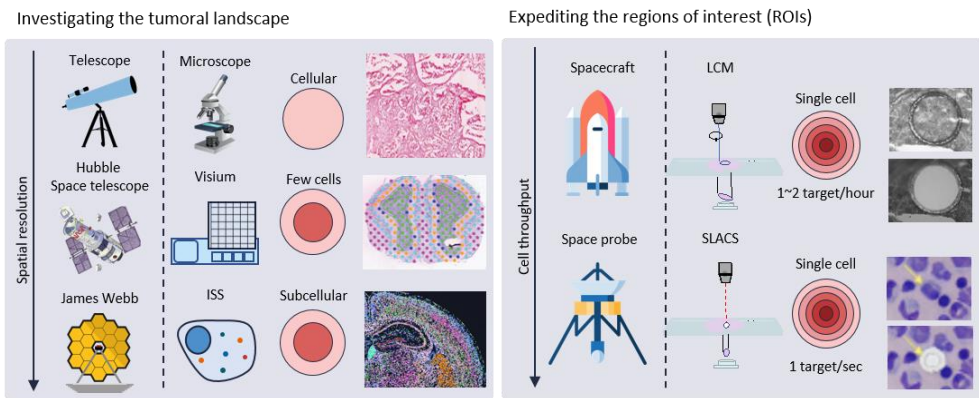


Figure 1.2 Technologies for investigating the tumoral lanscape and technologies for expefiting the region of interests.

Recently, a number of spatial omics technologies have been developed recently to address the questions in cell spatial biology. Just as telescopes have been developed to delineate the universal landscape, there are spatial omics profiling technologies developed to delineate the spatial landscape of the tissues [31]. On the other hand, there are technologies developed to characterize the specific regions of interest, just as spacecraft dive into specific planet for in-depth characterizations.

1.2. Spatial Landscape Profiling Technologies

1.2.1. Spatial transcriptomics profiling technologies

Spatial transcriptomics technologies was recognized by Nature Methods journal as the ‘Method of the year’ in 2020 [31], [32]. Thinking of the fact that bulk sequencing technologies have been selected as the ‘Method of year’ in 2005, and single cell sequencing in 2013, the spatial transcriptomics technologies are thought to bring a critical impact in cellular biology [33]. Currently developed spatial transcriptomics technologies have been considered as two common approaches, fluorescently tagged microscopic approach and in situ barcoded next generation sequencing approach [32]- [35].

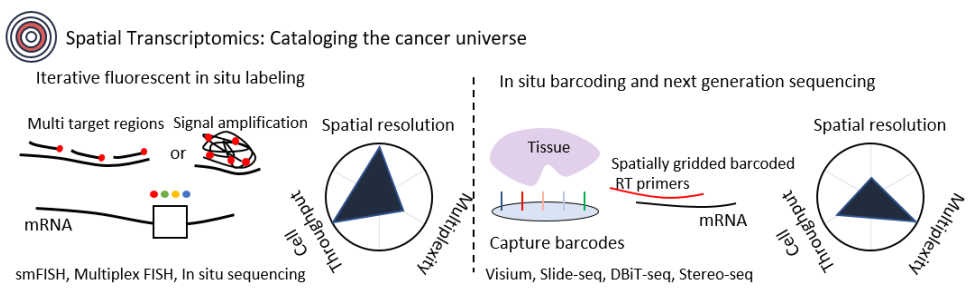


Figure 1.3 Spatial transcriptomics technologies for cataloging the cancer universe.

As addressed in the history of cancer research, microscopic approach is capable of observing the overall landscape of the tissue. By labelling the genetic molecules by fluorescence directly on the expressed position of cells, FISH is a historically

long-used technique to image transcripts [11], [13], [36]–[38]. It uses the fluorescence labeled target probes to detect the target transcripts. However, due to the spectrum overlap of the fluorescence molecules, there has been limitation in number of co-detectable transcripts. Various technologies have been attempted to increase the number of co-detectable targets with higher detection accuracy.

Single-molecule FISH (smFISH) used single capture probe to target the certain transcript, but further the method to target the different regions of the same transcript was developed to increase the signal-to-noise ratio [39], [40]. To effectively quantify the expressions of the transcripts, signal amplifiers were introduced in FISH methods [41], [42]. For instance, RNA scope implements the Z shaped capture probe to capture the target transcripts and labels the fluorescence probed to the capture probes, highly increasing the signal intensity [43]. FISH method has an intrinsic limitations, but due to its simplicity and specificity, FISH has been generally used to discover the local expressions of the target transcripts [44]– [48].

The iterative fluorescence approach to effectively increase the multiplicity of co-detectable targets. SeqFISH and MERFISH is the representative methods of iterative FISH. SeqFISH repeats the fluorescence hybridization and de-hybridization steps to identify the target probes [49]. It had to design the multiple fluorescently labelled target capture probes for each genes to increase the signal intensity. The process was highly time consuming and expensive, so they developed the SeqFISH+ method which introduced the signal amplifier concept increasing the signal intensity. SeqFISH is now commercialized by Spatial Genomics [50]. MERFISH also uses the

iterative methods to detect the signals, but instead they uses the signal amplifier to effectively identify the transcripts and adopted the error correction scheme which might occur during the iteration step [51], [52]. MERFISH is commercialized by Vizgen.

In situ sequencing methods were suggested which can sequence the genetic regions *in situ* [53]. Padlock barcodes were designed to target the transcripts for *in situ* amplifications by rolling circle amplifications (RCA) of the sequences and repeated the sequencing steps of the barcodes [54], [55]. In situ sequencing approaches are currently commercialized as Xenium. It has been applied to analyze the tumoral heterogeneity in breast cancer tissues [53], [56]. It minimized the design of the fluorescently labelled capture probes, reducing the cost of the assay. Similar concept were also developed as FISSEQ and STARmap which can directly sequence the targets on the tissue [57],[58], [59],[60].

In situ capturing of the transcripts in the spatially localized regions have been suggested enabling the *de novo* detection of the expressed transcripts. Spatially barcoded poly(dT) capture probes capture the Poly A tail of the mRNAs and are retrieved into the solution phase [61]. Then after next generation sequencing, the transcripts are deconvoluted as the spatial regions. Such unbiased way for mapping the transcripts enables the patterning of the cell types in a spatial landscape [62], [63]. The method has been commercialized as Visium by 10X genomics. It is still developing its technologies to increase the resolution and information depth [62], [63],[64],[65]. The technology have been actively developed to increase the

information depth of the transcriptome it can provide [66]–[69]. Additional microfluidic channel based *in situ* barcoding approach has been conducted as DBiT-seq. Such barcoding approach has limitation that only the 3' end of the transcript can be sequenced.

1.2.2. Spatial Omics Technologies

Currently, spatial transcriptomics is the technology that has been recognized a lot since transcriptomes can indirectly infer the snapshot of the cell status. Observing the spatial landscape of the transcriptomes infers the overall cellular outlines of the tissue, but additional information are required to understand the cell dynamics. Proteome are the functional unit of the cells and genome are described as the blueprint of the cells. Epigenome are responsible for regulating the genome for encoding the transcriptome. Therefore, to fully understand the cellular dynamics, multi-dimensional integration of the omics data are required for understanding the cellular functions. When compared to spatial transcriptomics, other omics profiling technologies are in its early states. Increasing number of researches are developed recently, and not many of those technologies has been commercialized yet.

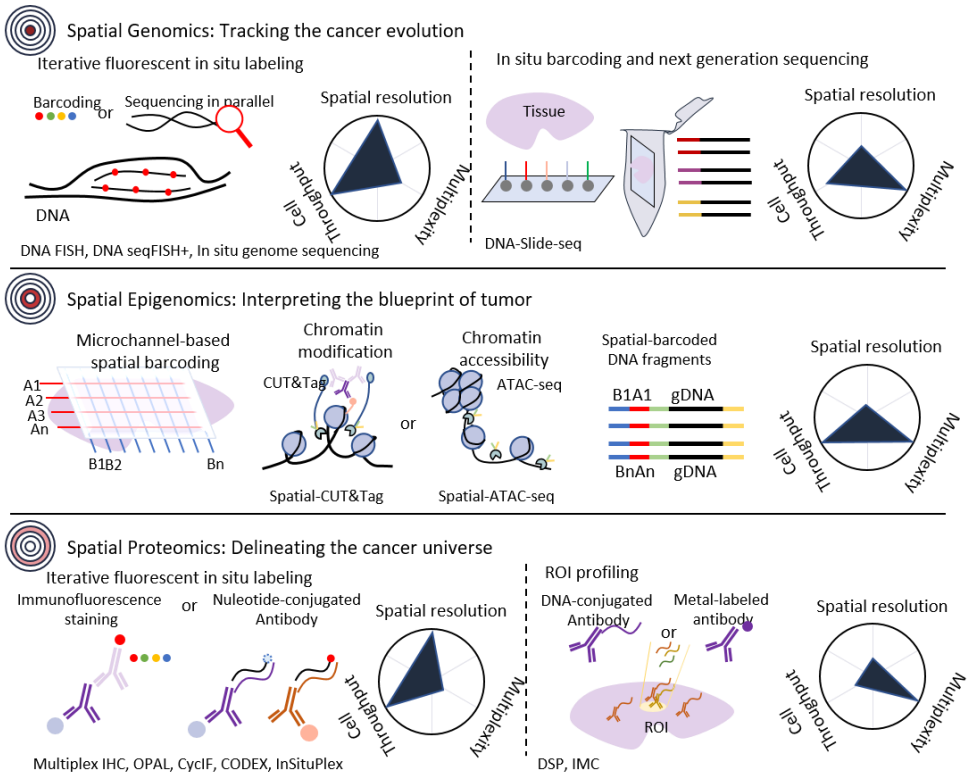


Figure 1.4 Spatial omics technologies for defining the onco-verse

Spatial genomics are important for defining the tumor progressing by analyzing the genomic aberrations of the cancer cells. FISH techniques has been also developed for targeting DNA sequences in tissue [70], [71]. *In situ* DNA FISH had been contributes in defining the cancer subtypes by defining single-nucleotide polymorphism [72], [73]. DNA FISH also shared the intrinsic limitations of the number of co-detectable genes due to the wavelength interference. To increase the number of the genes, various sequential FISH methods has been suggested recently. DNA seqFISH+ [74]. *In situ* genome sequencing (IGS) has introduced the iterative

imaging of the fluorescence labels and Slide-DNA-seq has adopted the *in situ* barcoding approach [75]. Until now, the genomic coverage of the listed technologies are not high enough to be directly applied to therapeutic target discoveries, but still spatial genomic profiling technologies are currently being applied to deciphering the tumor evolution, and spatial clonal populations.

Spatial epigenomics profiling technologies are in its very early stage. Since a number of epigenome assay chemistries have been developed recently, the approach to spatially index the spatial location via *in situ* barcoding and then apply the available epigenome chemistries are being developed [76], [77]. It was built on the microfluidic channel based in situ barcoding developed by DBiT-seq. The set of barcodes are delivered to the spatial region in perpendicular rounds, generating the grid of barcodes. Cut & Tag and ATAC-seq has been applied after the barcoding rounds to decipher the chromatin modification and chromatic accessibility [78], [79].

Proteins are the functional unit of the cellular mechanisms, and it has been developed for a long time to spatially map the protein expressions. As addressed in the history of the cancer biology, H&E staining and immunostaining methods have been detecting the surface marker proteins, helping researchers to draw the cell outlines or distinguish the cell types. To increase the number of protein targets, iterative methods for immunostaining have been developed to overcome the fluorescence overlap issue [80]–[82]. During the iterative process, cells go through the harsh antigen retrieval steps. Therefore, there is a limitation in the number of the detectable proteins. Gentler methods to strip off the previous fluorophores have been

developed by bleaching the fluorophore or introducing the intermediate reporter probes. The co-detection by indexing method (CODEX) uses the DNA-conjugated antibody and detects the antibody by fluorescence labelled DNA probes [83], [84]. By iterative fluorescence detection process of the DNA-conjugated antibodies, it has highly increased the number of the protein targets up to 60 targets. Spatial protein signatures reflects the cellular dynamics of the tissue, further stacking the understandings in cancer biology [85].

Furthermore, multi-dimensional integration of the omics information are applied for better understandings of the cancer cell biology [86]. Co-detection of the transcripts and the target proteins are relatively accessible when introducing the DNA-conjugated antibodies. DBiT-seq, a microfluidic channel based *in situ* barcoding approach is capable of indexing the transcripts with the protein targets by introducing the DNA-conjugated-antibodies [87]. Also, CITE-seq, a chemistry developed to analyze the target proteins and whole transcriptome sequencing, has also been spatially integrated using the microfluidic *in situ* barcoding platform [88]. Nanostring has developed the iterative FISH technique which can co-detect the transcripts with the proteins [89].

SM-Omics platform suggest the automated integration of the multiple single-omics profiling technologies [90]. It has integrated the protein measurement techniques using the DNA conjugated antibody or immunostaining, with the conventionally available spatial transcriptomics profiling technologies as well as the cellular phenotypes such as H&E staining image. Increasing number of researches

suggests the method of integrating two or more dimensions of omics information.

Spatial landscape technologies have advantage in that it can be applied to a large-scale patterning of the omics profiles. It has been contributed in the novel spatial findings in oncology in terms of spatial heterogeneity, cell types consisting the tumor microenvironment, and gene counts of the overall tissue. However, most of the spatial landscape technologies show low information depth. For instance, iterative FISH and *in situ* sequencing methods need the targeting of the transcripts, therefore, only the gene counts can be yielded. Also, *in situ* capturing technologies captures only the 3' or 5' end of the transcripts, resulting in the information of only the gene counts or the sequence of the 3' or 5' end region. Therefore, in current state, it is more focused on profiling the overall patterns rather than suggesting the specific therapeutics or diagnostics target biomarkers.

Table 1.1 Tumor research driven by spatial landscape technologies

Target	Technology	Findings Types	Applications in Cancer
RNA	Visium	Biomarker	Prostate cancer (Erickson et al. (2022)[69], Tuong et al. (2021)[91]), Melanoma (Hunter et al. (2021)[92]), DCIS (Nagasawa et al. (2021)[93], Bladder Cancer (Gouin III et al. (2021)[94])
	Visium	Heterogeneity	Prostate cancer (Joseph et al. (2021)[95]) Breast cancer (Andersson et al. (2021)[64]), Carcinoma (Luca et al. (2021)[96]), Pancreatic cancer (Sun et al. (2021)[97], Ma et al. (2022)[98])
	Visium	TME	Carcinoma (Ji et al. (2020)[99]), Breast cancer (Wu et al. (2021)[100]), Colorectal cancer (Wu et al. (2022)[22], Qi et al. (2022)[101]), Lung cancer (Dhainaut et al. (2022)[102]),

			Glioblastoma (Ravi et al. (2022)[103]), Neuroblastoma (Van de Velde et al. (2021)[104]), Mixed (Nieto et al. (2021)[105])
	Visium	Integrated	Melanoma (Thrane etal. (2018)[106]), DCIS (Wei et al. (2022)[107]), Pancreatic cancer (Moncada et al. (2020)[108])
	MERFISH	TME	Glioblastoma (Hara et al. (2021)[109]), Carcinoma(Magen et al. (2022)[110])
	RNAscope	Biomarker	Gastric cancer(Tamma et al. (2018)[111])
	RNAscope	Heterogeneity	Breast cancer(Annaratone et al. (2017)[47]), Cancer cell lines (Rowland et al. (2019)[48])
	ISS	Biomarker	Breast cancer (Ke et al. (2013)[53]), Glioblastoma (Ruiz-Moreno et al. (2022)[112])
	ISS	Heterogeneity	Breast cancer (Svedlund et al. (2019)[56])
	Fisseq	Biomarker	Breast cancer (Alon et al. (2021)[113])
Protein	Nanostring	Biomarker	Medulloblastoma (Zagozewski et al. (2022)[114]),
	Nanostring	TME	Carcinoma (Sharma et al. (2020)[115]), Colorectal cancer (Pelka et al. (2021)[116]) , Melanoma (Van Krimpen et al. (2022)[117]), Lung cancer (Wong-Rolle et al. (2022)[118])
	Nanostring	Integrated	PDAC (Han et al. (2022)[119]), Carcinoma (Sadeghirad et al. (2022)[120]), Prostate cancer (Brady et al. (2021)[121]), Breast cancer(Kulasinghe et al. (2022)[122], McNamara et al. (2021)[123]), Oral cancer (Schmitd et al. (2022)[124])
	mIHC	TME	Colorectal Cancer (Zhang et al. (2020)[125], Che et al. (2021)[126]), Glioblastoma (Pombo Antunes et al. (2021)[127])
	CODEX	TME	Colorectal Cancer (Schürch et al. (2020)[128])
	CODEX	Integrated	Lymphoma (Mondello et al. (2021)[129], Phillips et al.

(2021)[130]), Bladder cancer (Gouin et al. (2021)[94])

1.3. Spatial ROI-profiling technology

1.3.1. Previous sorting systems for spatial omics

There are spatial omics profiling technologies that focuses on in-depth characterizations of a specific region of interest, just as spacecraft focuses on getting deeper insight of the specific celestial bodies. Needle biopsy was initially used to biopsy the specific cancer regions, and proceed the post-processing chemistries [131], [132]. Emmert-Buck suggested the first Laser Capture Microdissection (LCM) which can dissect out the specific region from the tumor tissue sections [133]. Currently, LCM has two commercialized technologies; one using the UV-laser to dissect out the tissue, and one using the IR-activated EVA polymer to stick out the regions of interest [134]. LCM has a distinct advantage in that it can isolate the tumor regions and integrate the existing molecular assays. In the 20th century to early 2000s, LCM was introduced for DNA genotyping, RNA-sequencing, and protein analysis. Currently, LCM is being applied in various cancer researches providing detailed insights towards cancer [135], [136].

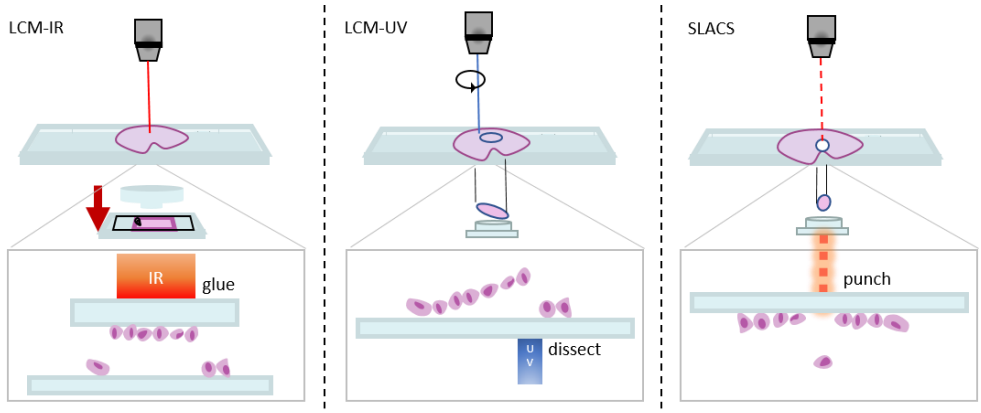


Figure 1.5 Spatial ROI profiling technologies for in-depth characterization

Unlike other spatial landscape technologies which provides low-depth *in situ* genomics analysis, LCM has been used to address the *de novo* genomic mutations such as single nucleotide variants (SNV) or copy number alterations (CNV) in heterogeneous tumoral regions or to trace the cancer evolution by tracking the genomic features of the subclones [137], [138]. Also, combined with RNA sequencing, *de novo* transcriptomic features were also deciphered in the spatial context. When combined with RNA-sequencing, isoforms or epitranscriptomic features can be thoroughly studied [139],[140]. Spatially differentially expressed biomarkers or pathway analysis of tumor progression or evolution has been studied through in-depth transcriptomic analysis. It was also introduced in proteomic analysis combined with the mass-spectrometry or other quantative profiling tools for protein analysis [141], [142]. Particularly, the biomarker discovery for diagnostics or therapeutics targets has been proceeded. The LCM-MS has a key advantage of *de*

novo identification of protein expression when compared to those that needs the design of specific antibodies to detect the expressed proteins. Also, since epigenomic research is in its early stage in the spatial omics field, ROI-profiling technologies has its advantage in that it can be applied to the cutting edge epigenomic assays [143], [144].

Table 1.2 Tumor research driven by spatial ROI profiling technologies

Target	Technology	Finding Type	Applications in Cancer
RNA	LCM	Biomarker	Prostate cancer (Cato et al.(2019)[145]), Breast cancer (Gómez-Cuadrado et al.(2022)[146]), Lung cancer (Chowdhuri et al. (2012)[147], Malapelle et al. (2011)[148])
	LCM	Heterogeneity	Nasopharyngeal cancer (Tay et al.(2022)[149])
	SLACS	Biomarker	Breast cancer (Lee et al.(2022))
DNA	LCM	Biomarker	Prostate cancer (Rubin et al. (2000)[150]), Bladder cancer (Cheng et al. (2001)[151], Cheng et al. (2004)[152]) Breast cancer (Bertheau et al. (2001)[153]), Carcinoma (Zhang et al. (2009)[154], Ellis et al. (2020)[138])
	LCM	Heterogeneity	Breast cancer (Wild et al. (2000)[155]), Carcinoma (Cheng et al. (2002)[156], Jones et al. (2005)[157]), Melanoma (Katona et al. (2007)[158]), Mixed (Olafsson et al.(2021)[159])
	SLACS	Heterogeneity	Breast cancer (Kim et al (2022))
Protein	LCM	Biomarker	Breast cancer (Cowherd et al.(2004)[160]), Ovarian cancer (Buckanovich et al. (2007)[161]), Glioblastoma (Lam et al.(2022)[162]), Carcinoma ()
Epigenome	LCM	Biomarker	Lung adenocarcinoma (Selamat et al. (2012)[163])
	LCM	Heterogeneity	Adrenocortical carcinoma (Schillebeeckx et al.(2013)[164])
	LCM	CTC	Lung cancer (Zhao et al.(2021)[165])
DNA + RNA	LCM	Biomarker	Breast cancer (Jovanovic et al. (2017)[166])
	LCM	Heteogeneity	Carcinoma (Chen et al. (2022)[167]), Lung cancer (Krysan et al. (2019)[168]), TNBC (Zhu et al. (2021)[169])

1.3.2. The need for the development of automated spatially-resolved cell sorter

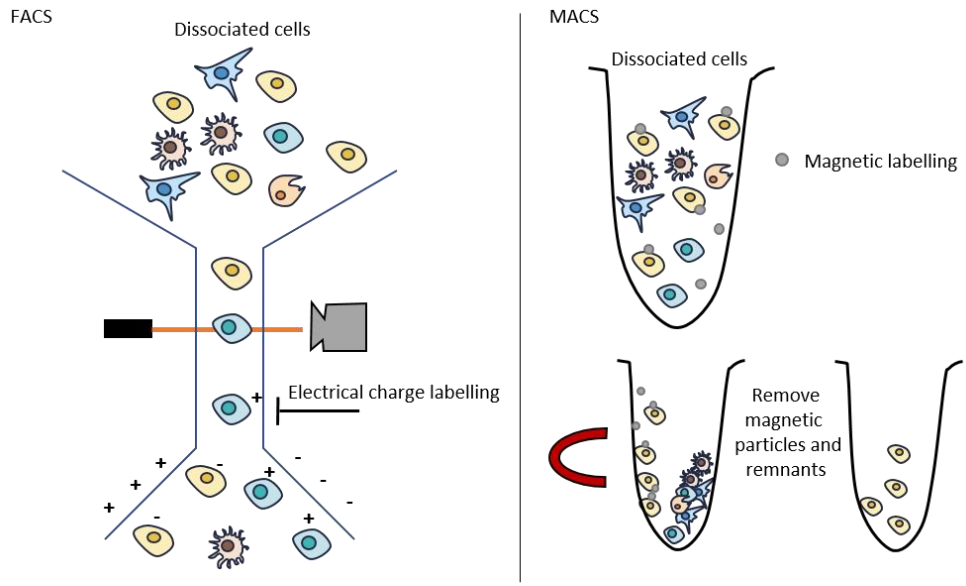


Figure 1.6 Conventional cell sorting technologies

Fluorescence activated cell sorting (FACS) and Magnetic activated cell sorting technologies (MACS) are mostly wide used cell sorting techniques in current state. FACS uses the fluorescence tags to target the cells of interest, and MACS uses the magnetic particles. FACS uses the microfluidic channel to flow the cell to the cytometry machine in a single cell level, then computes the fluorescent information and tags the corresponding electrical charge label to sort the cells to the reservoir. The multiplexity of the single sorting process of the conventional FACS is up to 3

and unless $10^5 \sim 10^6$ cells are needed for the sorting process. In case of MACS, magnetic particles are added to tag the cells of interest, then pulls the cells of interest via magnetic interactions. Then remnant cells are washed out, keeping the tagged cells to the desired reservoir. Another novel cell sorting technologies were developed to sort out the cells of interest by using the cellular phenotype information using Raman imaging and machine learning [170], [171]. However, the listed technologies all need the dissociation of the cells to the solution phase to sort out the cells via microfluidic flow or magnetic interactions, therefore, losing the spatial context of the tissue. Still, they were recognized by its throughput and applicability to the post-processing chemistries and have contributed in thriving the single cell sequencing field.

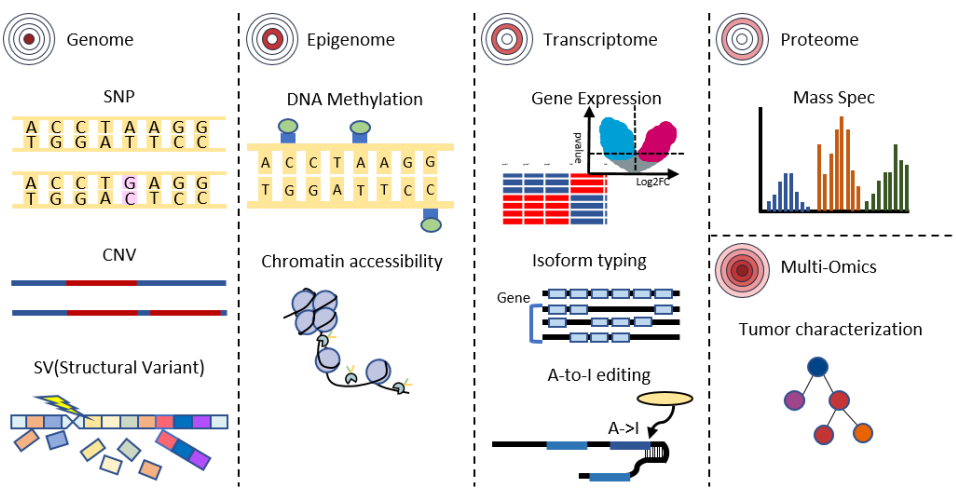


Figure 1.7 Appications to existing molecular profiling technologies

Conventional spatially resolved cell sorters introduced in section 1.3.1 spatial sorting systems such as LCM and SLACS have the capability of sorting out the region of interest and apply the listed post processing chemistries, but it still has unsolved problems to be generally applied to the life scientists. Conventional LCM device has low throughput that it takes about an hour to sort two regions. Also, the devices used for the dissection such as UV or the EVA polymer might induce the cell damage during the dissection step. For instance, it is applicable to pull out the single or few region of interest not applicable for to sort out the tumor infiltrated immune cells, but not for the hundreds of tumor infiltrated immune cells that might provide the therapeutic target sequence. As single cell analysis has thrived for last ten years for its novel discoveries in life science due to the high-throughput droplet based single cell sequencing technologies, spatial cell sorting technologies should also be developed for in-depth sequencing of the target cells as well as the spatial landscape technologies.

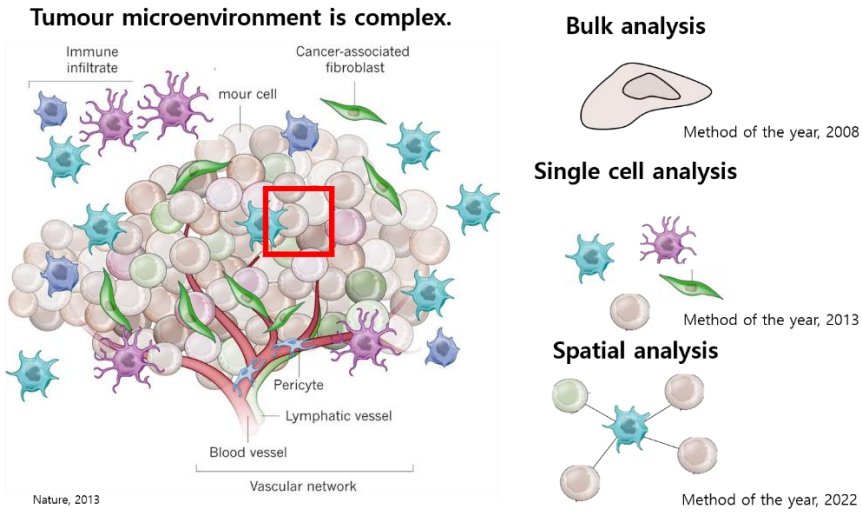


Figure 1.8 The era of spatial analysis will lead the active research in tumor research

1.4. Outline of the dissertation

1.4.1. Previous work from our group

To address the throughput and cell damage issues of the conventional LCMs, spatially resolved laser activated cell sorter (SLACS) have been developed from our lab [172]. SLACS utilizes the indium tin oxide slide glass for tissue to be located, then the IR activates the evaporation of the target regions. Compared to other LCM devices that use UV laser or EVA polymer to dissect out the specific regions, SLACS uses the laser-activated evaporation, so it takes about target per second time to isolate the desired region. The size of the isolation differs from 1 μm diameter to 1mm

which can vary by the user's need. Conventional staining methods such as H&E, immunostaining, FISH, etc. can be applied to target the regions of interest and also the other spatial landscape technologies can be applied to guide the cells or interest. Retrieved cells can go through the post-processing such as NGS-based assays or mass spectrometry assays. SLACS has conducted various researches in cancer research such as epitranscriptomic analysis in spatially located cancer stem cell-like microniches [173]. Novel adenosine to inosine A-to-I edited GPX4 gene has been discovered as the triple negative breast cancer (TNBC) clinical outcome prediction marker. Also, full length sequencing led to the profiling of the B cell receptor and T cell receptor sequences for specifying the therapeutically actionable immune sequences.

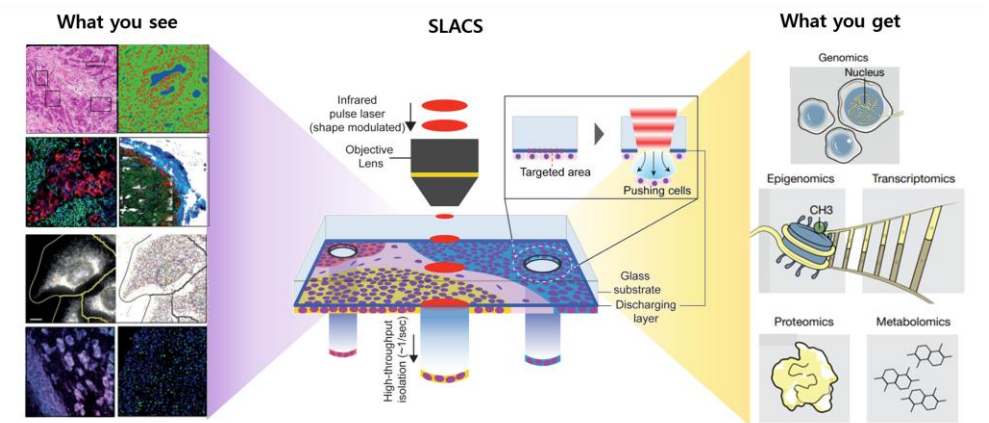


Figure 1.9 SLACS: Connecting spatial and molecular information

The first demonstration of the SLACS device was reported in Spatial DNA

sequencing in breast cancer tissue sections. By analyzing different microniches using multiple displacement amplifications (MDA), the subclonality and the evolutionary relationship between different subclones were revealed [172]. Also, by using the serial sectioned tissues, 3D genomic map showing the different subclones could be analyzed with whole genome sequencing, whole exome sequencing and targeted sequencing. SLACS was able to construct the 3D genomic landscape of the breast cancer, providing intra-tumoral subclones with their genomic aberrations such as copy number alterations, single nucleotide variants and structural variations for tracking the tumor evolution. The discovery could give the insight towards personalized diagnostic tools for circulating tumor DNA (ctDNA) panels. Also, SLACS was applied in isolating the circulating tumor cell (CTC) in CTC capturing biochips, where the CTC was captured in the EpCAM coated pillar structures [174].

SLACS was applied to the post processing of analyzing the spatial transcriptomics and epitranscriptomics [173]. Unlike other spatial landscape technologies, SLACS was able to analyze the full length sequencing of the transcripts. Two different cancer stem cell markers ALDH1 and CD44 was used to label the different cell populations. The differently labelled microniches had a unique adenosine to inosine (A-to-I) editome sequences. They further revealed that A-to-I edited GPX4 had a high correlation to the prognosis of the triple negative breast cancer patients with neoadjuvant chemotherapy. Also, full-length immune cell receptor sequences could be profiled, proving the potential to be applied to the interactive molecular analysis. For instance, tumor infiltrating immune cells in the

tumor microenvironment can have the specific antibody sequence that can be developed for the diagnostics or therapeutics targets. Aside from the DNA or RNA sequencing, mass spectrometry or epigenomic assays also can be applied after the isolation, further expanding its applicability towards the spatial omics assay.

1.4.2. Main concept: Automated cell sorting system for multi omics

In this dissertation, I propose the novel technological solution for the spatial ROI-based omics profiling field. The main concept of the proposed technology is to automatically targeting the regions of interest and sorting ROIs for omics applications. Proposed method will address the current limitation of the SLACS device that targets the regions of interest manually. I suggest the automated What You See Is What You Get (WYSIWYG) cell sorting platform that enables the detection of the phenotypically marked cells to be isolated and connection to the further post-processing chemistries for the future spatial omics applications. Like the conventional FACS system which can automatically isolate the specific targets tagged with the fluorescent labels, I suggest the automatic spatially-resolved laser-activated cell sorting platform for the ease-of-use to the general life scientists.

1.4.3. Outline of the dissertation

In this dissertation, the automated cell sorting using the SLACS device is described. In Chapter 2, the assay quality of the retrieved cells using SLACS is described. In Chapter 3, development of the automated targeting and isolation of the

cells and applications are described. Finally, in Chapter 4, the summary of the dissertation and future direction using the device are described.

.

Chapter 2.

Development of the spatially resolved laser-activated cell sorting platform

In this chapter, I will describe the development of the SLACS device and its performance. SLACS device is capable of isolating the regions of interest of the pre-marked regions from the image samples. SLACS has its advantage in that it can be applied to various post-processing chemistries including whole genome sequencing (WGS), full-length RNA sequencing, targeted sequencing and so on. By isolating the regions of interest, SLACS has its capability of providing the in-depth molecular information of the target regions.

2.1. Development of the SLACS system

2.1.1. Advantage of the SLACS system compared to other cell sorting technologies

Significant advances in imaging technologies and molecular cellular biology have brought biological discoveries that boosted the understanding of biological phenomenon. Ranging from hematoxylin and eosin (H&E) staining to immunofluorescent staining (IF), different staining modalities have laid foundations for modern medicine, and next generation sequencing (NGS) caused a quantum leap in biological discoveries in a decade [175]. While the imaging technologies provide information of the biological circuitry by providing spatial and structural information of a biological system, molecular cellular biology technologies provide the state of the biological cells in a more microscopic manner. However, to fully understand how the cells are functioning within spatial context, there needs to be a tool to connect the imaging technologies to molecular cellular biology [176]. Such integrated technologies that can sort out cells with preserved spatial information will connect the data from spatial assays to that from the molecular assays.

Conventional cell sorting technologies, however, are mostly based on microfluidic cell sorting. Fluorescence activated cell sorter (FACS) is perhaps the most widely used cell sorter [177]. FACS utilizes laser source to sort cells with the same fluorescence spectrum. These technologies require the cells to be dissociated into solution phase, losing the spatial information during the process. Intelligent

image-activated cell sorting and Raman image activated cell sorting measured the cell image with machine learning and Raman image of the cells, respectively [170], [171]. These techniques are able to preserve the cellular phenotype information because the cells are sorted according to their images, but these cell sorting techniques still require dissociation of the cells and therefore the spatial context is lost before the cells are sorted. Spatially targeting cell isolating devices such as laser capture microdissection (LCM) or similar methods are able to dissect out the regions of interest. Some platforms offer high resolution cell microdissection down to single cell level [134], [178]. However, the throughput for these platforms is three target retrievals per run (which takes approximately one hour), hindering their usage in cell sorting.

SLACS preserves the spatial information and demonstrate its usage in bridging spatial assay to various molecular assays, which isolate cells on an indium tin oxide (ITO) coated slide glass into commercialized PCR tubes based on their image. The properties of cell sorting such as accuracy are demonstrated and the possibilities of post-processing such as DNA sequencing, RNA sequencing, or other molecular biology techniques are validated, proving the potential of SLACS for novel biological discovery tool by sorting mammalian cells to bacteria. Especially, studies have reported a method of sequencing the genomes of small number of cells or single cell through this methodology [173], [179].

2.1.2. Workflow and design of the SLACS system

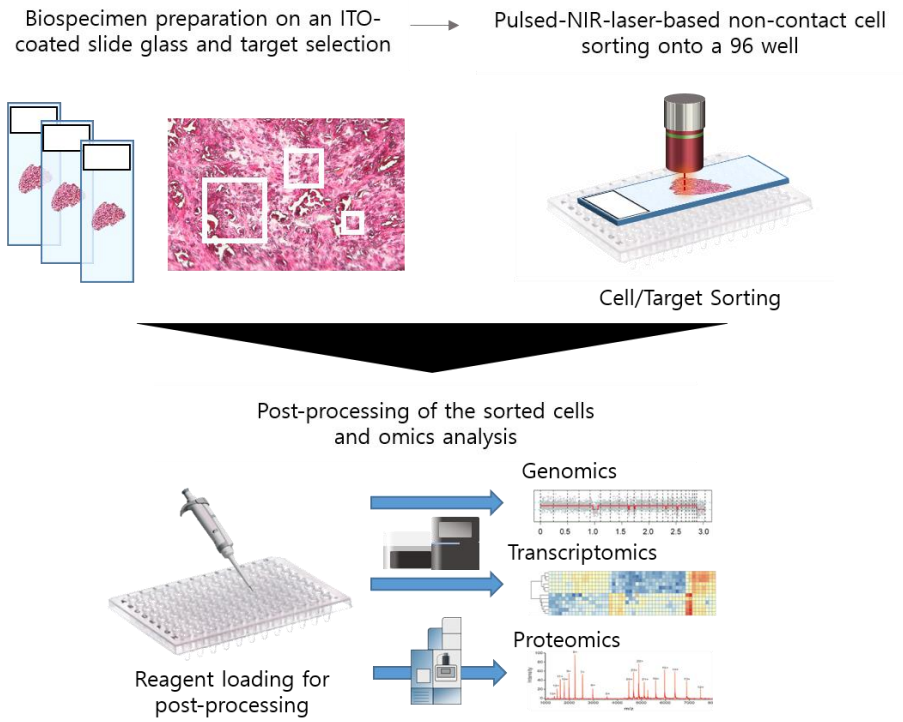


Figure 2.1 The work flow of the SLACS platform

SLACS works with a sample that is prepared on an ITO coated slide glass. Here, ITO layer serves as a discharging layer that readily vaporizes when exposed to the near-infrared (NIR) laser with a wavelength of 1064 nm in the cell sorting step, allowing cells on the exposed region to be sorted into the commercialized PCR cap on the retrieval stage of SLACS device. After biospecimen is prepared on an ITO coated slide glass, various staining methods with desired modality can be applied to the biospecimen without affecting cell sorting process, and in this report we used

Giemsa banding, Wright-Giemsa staining, H&E staining, immunohistochemistry (IHC) staining, and gram staining as representative staining methods.

Then, the targets are selected and isolated based on the real-time image of samples acquired as a CCD camera of SLACS instrument or the whole-slide image took by a whole-slide scanner or a general fluorescence microscope, preserving the spatial and the staining information of isolated cells. If a reference picture is generated, the target marking software can be used for supervised target selection. If the automated marking program is developed, both supervised and unsupervised cell sorting are available according to the demand for target selection from images.

In optical modules, real-time images of samples are acquired like microscopy with a CCD camera, and fluorescence light source and filter cube array with three filters is also embedded within the device in order to increase the range of stained samples. A reflective bright field light source is installed so that the target area can be lighted and seen on the CCD camera. In addition, the NIR laser from an Nd:YAG nanosecond pulsed laser source (λ = 1064 nm, pulse = 6 nanosecond) is used for ablating the target through the fixed light path. We chose 1064 nm pulsed laser because the laser with higher wavelengths than ultraviolet lasers is widely used for manipulating intact cells such as in optical tweezers, cell sorting methods, and other applications.

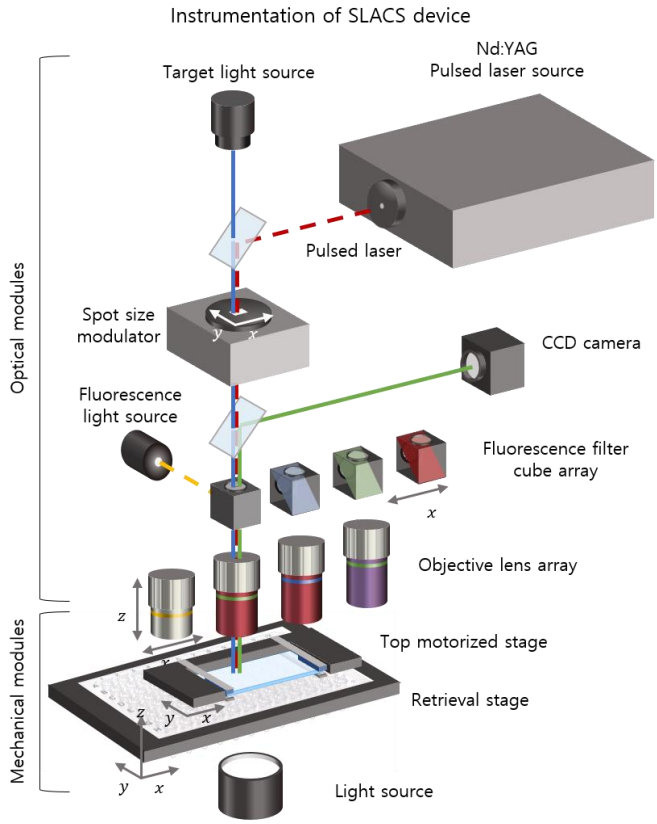


Figure 2.2 The optical and mechanical module of the SLACS device (Reprinted from [173])

Its spot size is controlled by the slit size adjusted by the x-axis and y-axis of the spot size modulator and the magnification of the objective lens. We used objective lens with long working distances that not only serves to focus the laser spots, but also provides enough working space to align the biospecimen to the retrieval plate. Next, mechanical modules comprise two motorized stages with sub-micrometer

scale precision: the sample stage and the retrieval stage. The top motorized stage to hold the sample-mounted ITO glass is controlled by z-axis to focus the real-time image, and by x-axis and y-axis to aim the target because the path of laser is fixed. On another motorized stage, called the retrieval stage, commercialized PCR tubes are arranged and placed to retrieve the isolated cells, By integrating the cell sorting method using ITO and NIR laser into the conventional microscopy, spatial information and molecular information of cells can be connected.

SLACS can be used for automated and manual cell sorting, as the in-house user interface with operating software and selecting target software designed for this SLACS device enables automatic and manual control of two motorized stages with sub-micrometer scale precision. Using the operating software and interface written with Python, the spot size modulator, the top motorized stage, the retrieval stage, and the objective lens are easily controlled manually like controlling the coordinates of the retrieval stage with alignment to the light path of the laser. The software panel is designed to move the retrieval stage so that the target can be retrieved onto the well number of interest in a 96-well. In addition a remotely target selecting software measures the coordinates and areas of the targets of interest, enabling cells to be automatically sorted based on this coordinates.

After the targets are selected remotely, the coordinates of the slide are taken to the SLACS instrument, of which the software converts the picture coordinates to the actual absolute coordinates of the sample. As the target is isolated, the device modulates the spot size matched to the target area and the top motorized stage is

moved so that the target is precisely below the objective lens. The target area was automatically calculated according to the objective lens and the selected area. With automatic run, each and every selected target is isolated in each well, and the maximum multiplexity per run is 96, which takes approximately one minute. If the well is switched to an empty well, the automatic run can be continued to isolate additional 96 targets

It was able to isolate single bacterium, single chromosome, single cell from bone marrow smear, cancer cell clusters from tissue sections. Especially, the samples prepared with different staining modalities to demonstrate the universality of the SLACS. The bacterium isolated was E. coli that was gram stained with S. aureus. The rod-shaped bacterium in contrast to the round-shaped S. aureus was sorted based on the bright field image.

Also, to show the high resolution applicability of the laser spot, we demonstrated single chromosome sorting from the Giemsa banded chromosomes. To our knowledge, this is the first report of isolating single chromosome using laser-based instruments. Then, the single cells from Wright-Giemsa stained bone marrow smear were sorted according to the cellular phenotype that became apparent by the staining. This is significant because while the golden standard for distinguishing cells from hematopoiesis is distinguishing the cells through staining, current advancements rely on predicting cell types through gene expression levels that can be contradicting to the traditional cell typing. Different tissue sections in breast cancer and bone tissue were demonstrated to demonstrate the capability of cell

isolation in soft and hard tissues. The size can vary from 1μm to 1mm range, which depends on the user's demand.

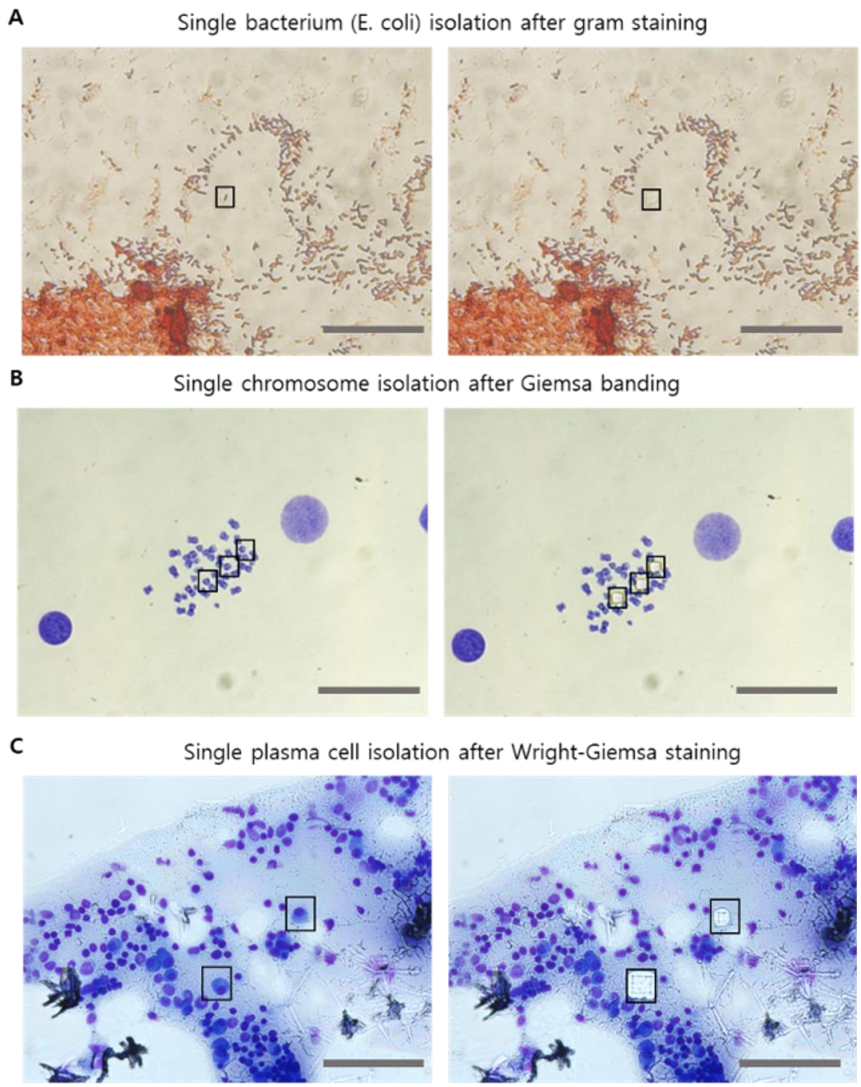


Figure 2.3 The isolation of the single biospecimen using SLACS device

(Bacterium, chromosome, cell)

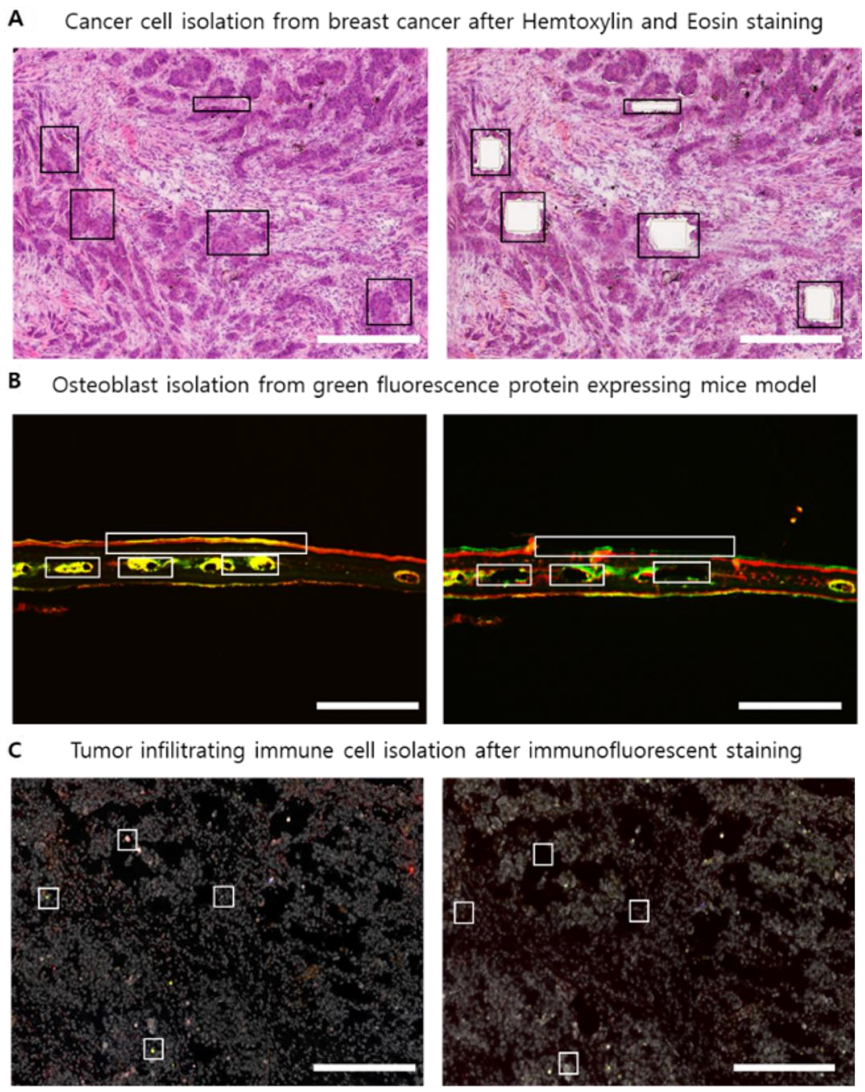


Figure 2.4 The isolation of the single biospecimen using SLACS device

2.2. The quality of the spatial omics assays after cell isolation via SLACS

There exists the tradeoff between the harshness of fixation or staining protocol and the quality of the biomolecules, but it was able to analyze different single nucleotide level sequencing analysis from the cells. For instance, the whole genome sequencing was performed using SLACS by utilizing the multiple displacement amplification methods followed by the proteinase K based cell lysis. The whole genome were amplified to perform the library preparation for whole genome sequencing.

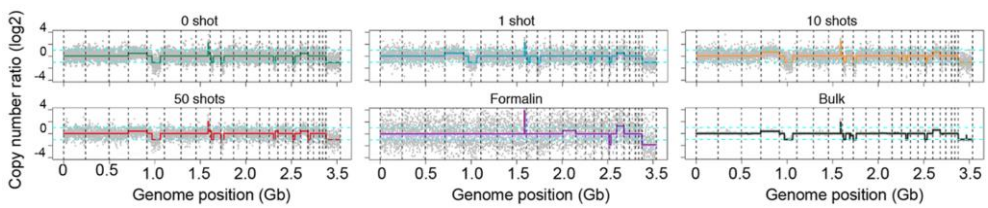


Figure 2.5 The CNV data after the irradiation of the IR laser (Reprinted from [179])

It was confirmed that the copy number variation data of the single HL-60 cell isolated using the SLACS showed the high correlation with the bulk sequencing data. Formalin fixation is known to generate the DNA fragmentation, so the formalin fixed data showed the mismatch with the bulk sequencing, but data of single cell isolated with pipette, single shot, 10 shot and 50 shots of the laser pulse during the SLACS

had the similar profile with the bulk data.

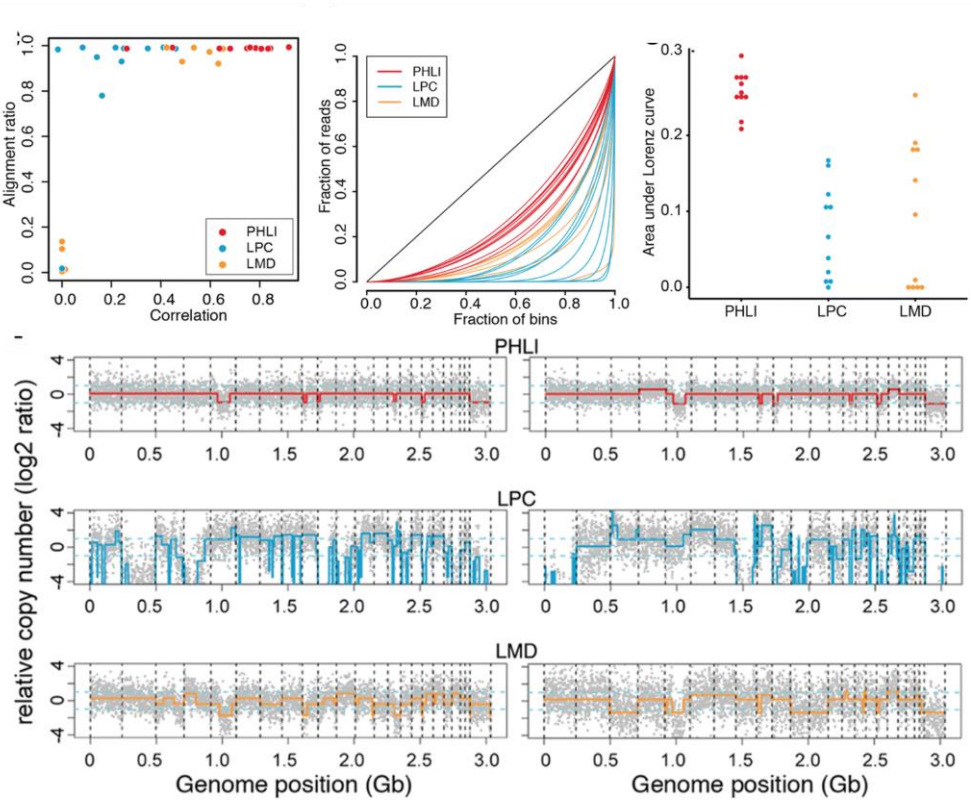


Figure 2.6 The comparison of the DNA quality to other laser microdissection methods (Reprinted from [179])

Also, it was confirmed that the quality of whole-genome sequencing data showed higher performance compared to other laser microdissection methods such as LPC or LMD platforms, which uses EVA polymer-based or UV laser-based isolation of the target regions. SLACS isolation of the target single cells showed better performance by evaluating the genome alignment ratio, which are portion of

the reads aligned to the human genome, and copy number correlation ratio, which are chromosomal copy number correlation with known HL-60 cell line profile. Also, Lorenz curve which shows the genome coverage analysis confirmed that cell isolation using SLACS showed highest genome quality.

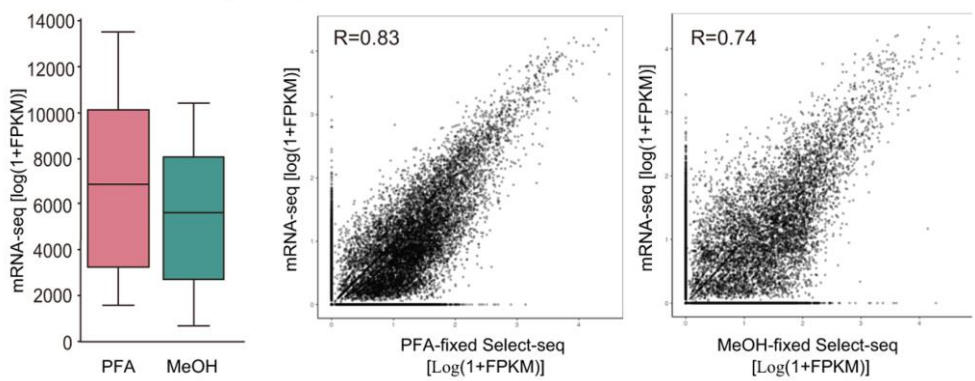
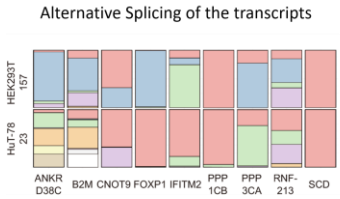


Figure 2.7 The quality of the full-length RNA sequencing after isolation using the SLACS (Reprinted from [173])

The performance of preserving the biomolecules were also confirmed in full length RNA sequencing data after the SLACS isolation. Fragment per kilobase of transcript per million mapped reads (FPKM) values were plotted in the SLACS isolated PFA-fixed and Methanol-fixed HEK 293T cell lines. The quality of the biomolecules differs according to the fixation method of the samples. When compared to the bulk mRNA-seq data, PFA and methanol fixation method showed the correlation of 0.83 and 0.74, respectively.



BCR sequence analyzed from the IM-9 cell line

Sample	Isotype	HCDR3	V gene	J gene	LCDR3	V gene	J gene
IM9_C4	G1	TKRRGVTDIDPFDI	IGHV3-9	IGHJ3	QHYNRPWT	IGKV1-5	IGKJ1
IM9_D1	G1	TKRRGVTDIDPFDI	IGHV3-9	IGHJ3	QHYNRPWT	IGKV1-5	IGKJ1
IM9_D5	G1	TKRRGVTDIDPFDI	IGHV3-9	IGHJ3	QHYNRPWT	IGKV1-5	IGKJ1
IM0_E3	G1	TKRRGVTDIDPFDI	IGHV3-9	IGHJ3	QHYNRPWT	IGKV1-5	IGKJ1
IM9_G5	G1	TKRRGVTDIDPFDI	IGHV3-9	IGHJ3	QHYNRPWT	IGKV1-5	IGKJ1

Figure 2.8 Transcriptome information that can be extracted from full-length RNA sequencing (Reprinted from [173])

It was able to profile the alternative splicing features or the heavy chain and light chained paired BCR sequence was able to be sequenced unlike other spatial transcriptomics profiling that captures the specific target regions of the transcripts or profiles only the 3’ or 5’ end of the transcript.

Chapter 3.

Automated image-based cell sorting platform

In this chapter, the development of the image based automated cell sorting platform is described. Current SLACS device needs the manual selection of the targets to be isolated. Therefore, it needs to identify the targets from the image and extract the target coordinates from the images. First of all, targets should be analyzed via image processing and it should be confirmed that the target position should be converted to the desired position at the SLACS device. Therefore, it is important to connect the image processing to SLACS system for automated cell sorting. The system was first validated the isolation of the desired targets using the encoded microparticles which are frequently used to mimic the micro-scale condition of cells. Then, it was applied to automatic isolation of the CTCs using the cell lines, and further utilized in clinical samples. Also aside from immunstainig methods, various

staining methods are applied to specify the target cell information such as *in situ* sequencing.

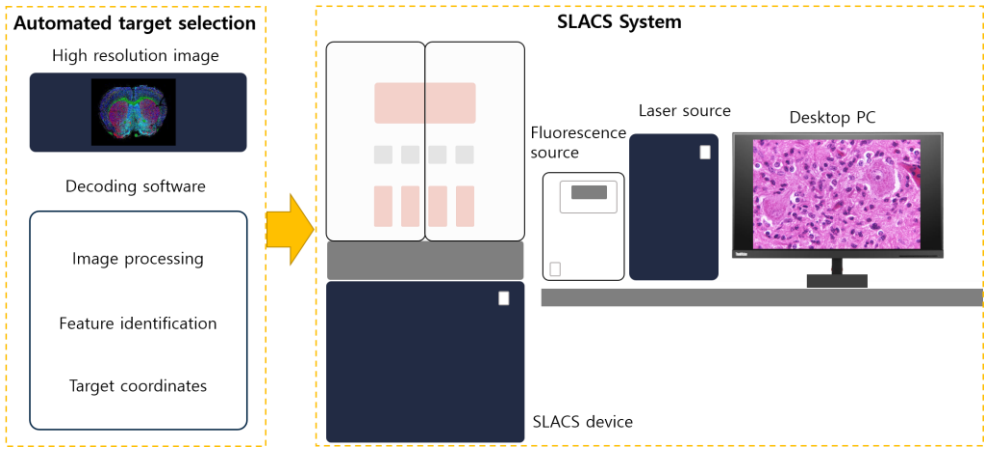


Figure 3.1 Automated spatially resolved laser activated cell sorting platform.

3.1. Validation of the automated targeting and transferring using the encoded microparticles

3.1.1. Design of the encoded microparticles for the validation

Encoded microparticles have the advantage that it can mimic the micro-scale condition of the cells, and the encodings can be precisely controlled [180]. Therefore, I used the encoded microparticles to validate connecting the automated target selection with the SLACS system to identify that the desired microparticles are transferred from the target positions.

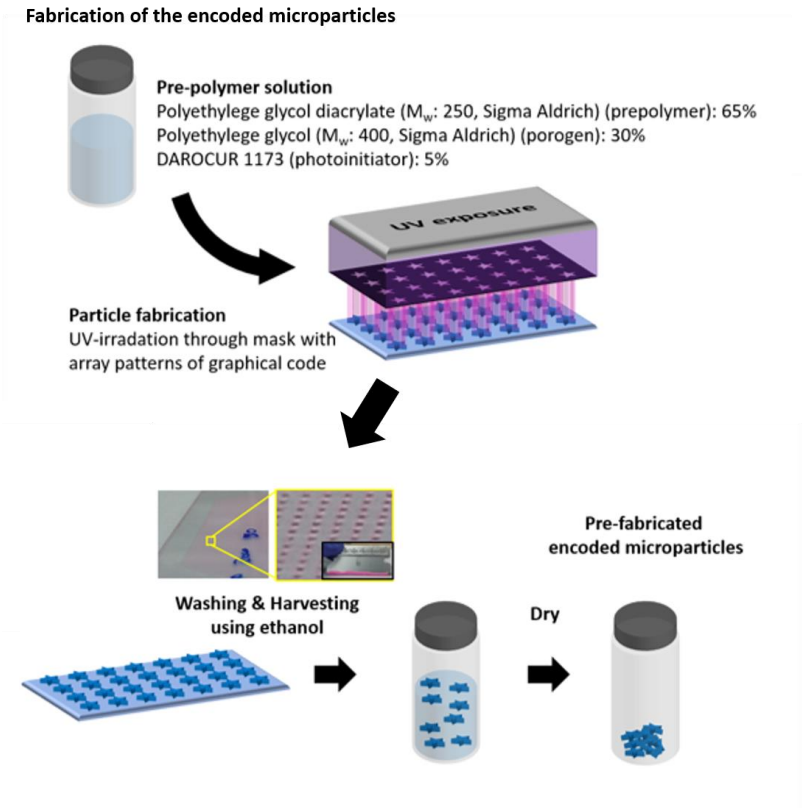


Figure 3.2 Fabrication of the encoded microparticles

Each encoded microparticles are first fabricated by the photolithography using the photomasks with homogeneous codes [181]. The homogeneous microparticles are generated when a mixture of pre-polymer and photoinitator is irradiated by the ultra violet laser beam. After the polymerization, uncured pre-polymer solutions are washed out and are harvested to the each separated tubes.

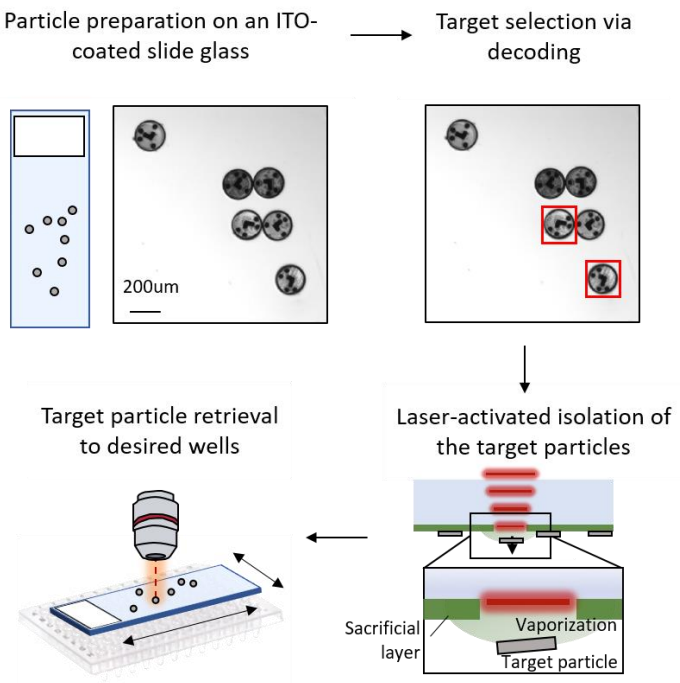


Figure 3.3 Schematic of the automated transfer of the target microparticles

To mimic the automated cell sorting condition, ten different encoded micoparticles were mixed and dispersed to the indium tin oxide (ITO)-coated glass slides, and the microparticles with the target codes were selected by decoding. Then the target coordinates were transferred using the SLACS system to the desired well coordinates.

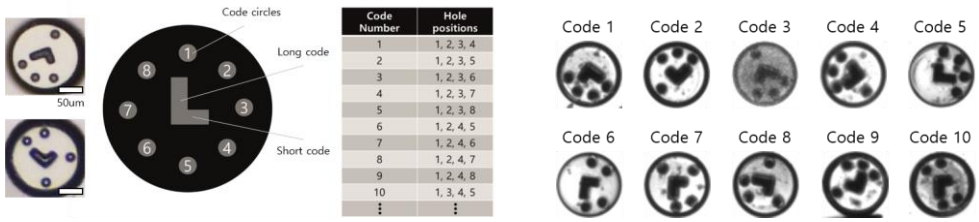


Figure 3.4 Code design for the microparticles

The code design for the encoded microparticles are listed in Figure 2.4. Long and short codes are designated for the alignment even in the rotated and flipped conditions. Each codes are fabricated using the photolithography which are capable of mass production.

3.1.2. Neural net-based pattern recognition for decoding the encoded microparticles

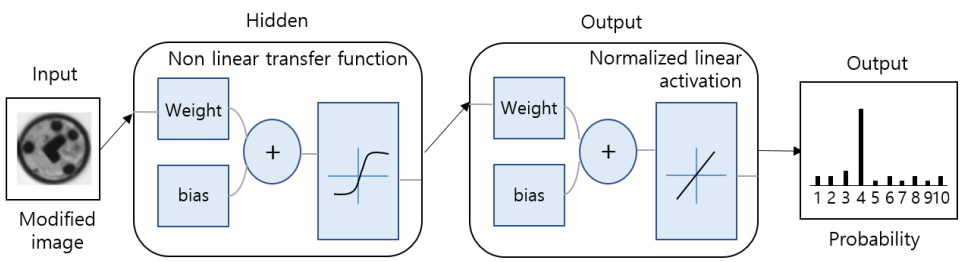


Figure 3.5 Neural net-based pattern recognition program

The decoding schematic is designed using the neural net-based pattern recognition program. The neural net-based pattern recognition programs are highly dependent to the image training set, so other approaches such as convolutional neural

network or graphical neural network is more adequate with the condition of cell or tissue. However, in the case of the encoded microparticles, it was easier to gain the homogeneous images of the single tags, therefore, was able to apply the neural network-based pattern recognition [181]. The modifies images and the labels are entered to the decoding program as an input and 70% of the images are used as a training sets, 15% for validation set and 15% for test sets. The output of the neural net is the probability of the expected labels.

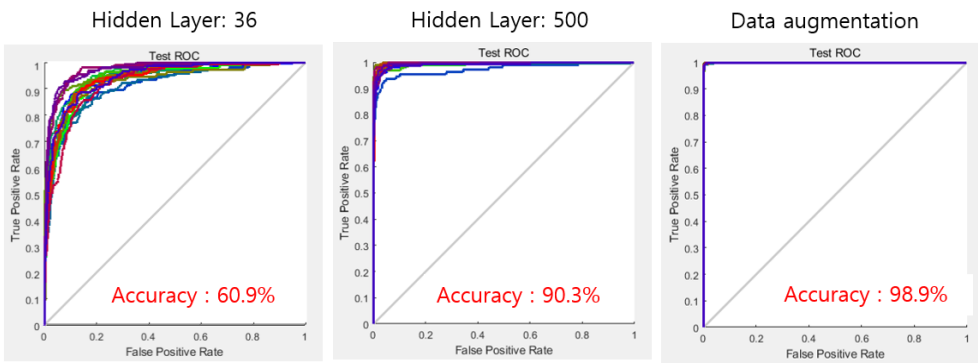


Figure 3.6 Decoding accuracy enhances as the number of hidden layer and training sets increases

The receiver operating characteristics (ROC) curve which is a graph that shows the performance of the classifications is described in Figure 2.6. The accuracy of the decoding increases as the number of hidden layers increase. Also, the data augmentation resulted in the enhanced accuracy.

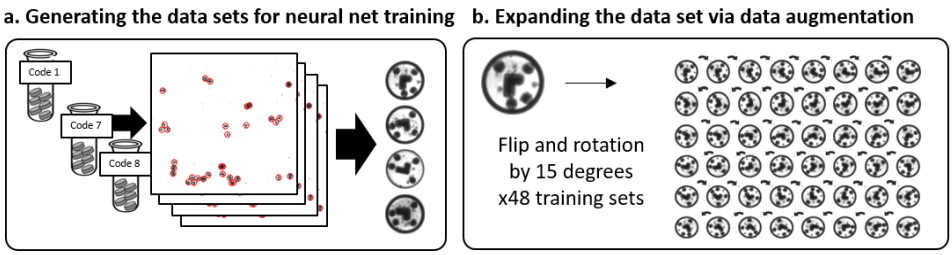


Figure 3.7 Generation of the training sets for the decoding

The training sets for training the neural net-based pattern recognition program was generated by extracting the images of the particles. After dispersing the homogeneous particles on the glass slide and extracting the particle images by circle detection. Then to increase the training sets virtually, data augmentation was proceeded by flipping and rotating each extracted code images.

The confusion matrix maps the misalignment of the actual and expected result of the classification. I trained ten codes of the microparticles and tested the performance of the decoding algorithm. The accuracy of the decoding of each ten codes were ~99%. The accuracy of the decoding can increase when increasing the number and quality of the data sets. Then, I ran the decoding algorithm to identify each codes from the dispersed images of the ten different microparticles and identified the microparticles by using the neural net based decoding algorithm.

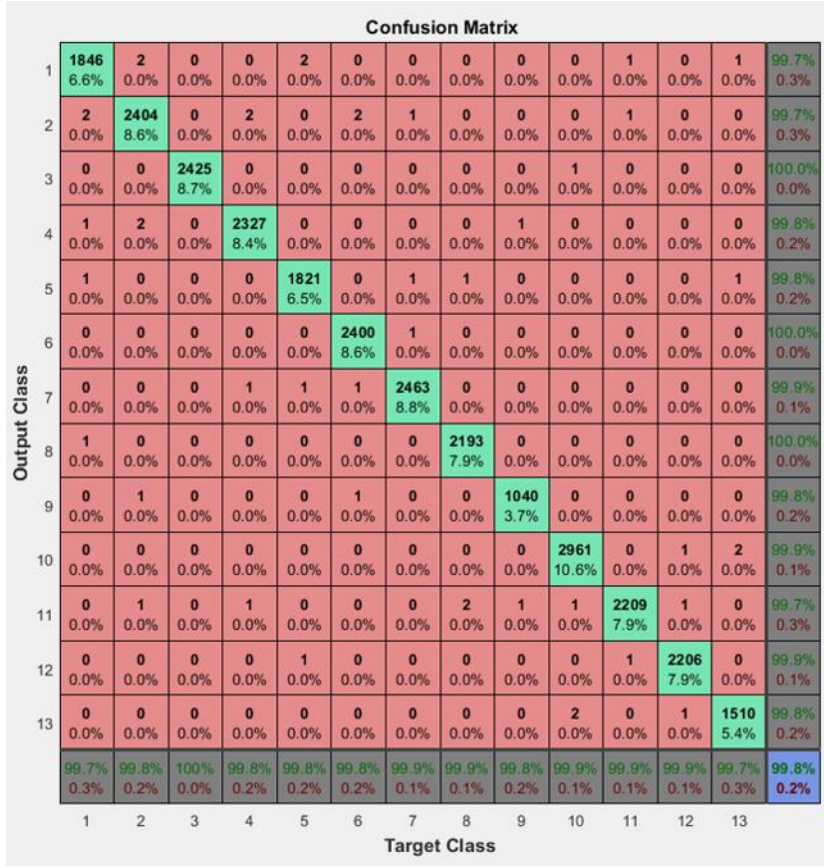


Figure 3.8 The confusion matrix of the decoding program

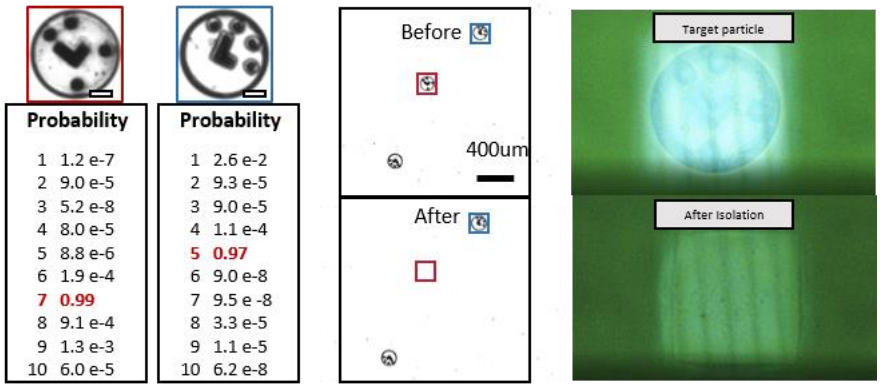


Figure 3.9 The extracted ralative coordinates of the target particles and result of the probability and isolation of the desired microparticles

After extracting the target coordinates of the relative coordinates in the image, it was possible to calculate the actual coordinates at the machine, using the reference markers for calculating the relative positions. As shown in the Figure 2.10, it was capable of classifying the encoded microparticles and connect it to the SLACS device. It was possible to isolate the desired targets using the laser isolation. Then, I tried to validate that retrieval of the images were able to perform at the designated well of the 96 well plates.

Figure 3.10 Transfer of the food dye loaded microparticles to the designated wells

3.2. Automated cell sorting for targeting the rare cells

3.2.1. The need for automated cell sorting in isolating the circulating tumor cells (CTCs)

Automated spatially resolved cell sorting contains its advantages in dealing with the rare cells, which are difficult to be applied to the FACS that needs at least 10^5 cells to run the isolation. CTCs are tumor cells that circulates in the blood which might be in the metastatic state of the tumor cells. CTCs are known and recognized to contain the information of the metastatic cascade and the characteristics of the primary tumors. Therefore, increasing researches are insisting that the CTC analysis might provide the information of the solid tumors which are difficult to collect from the real patient.

However, CTCs are difficult to detect from the massive blood cells, since CTCs exist in very low-abundance in blood. Therefore, a number of approaches have developed to purify the CTCs from the blood plasma based on its physical properties or biological properties. CTCs has a size range of 12 to 30μm, therefore is generally bigger than other cells. Therefore, by using the microfluidic approaches to isolate the cells according to its size, it is possible to purify the CTCs, but since white blood cells (WBCs) have similar size range with average of 8 μm, it is difficult to isolate only the CTCs from the blood plasma [182]. Also, the biological methods to isolate CTCs only are capable of isolating CTCs by capturing with certain markers such as EpCAM [183]. Therefore, numerous approaches are currently being developed to

effectively isolate the CTCs from the blood samples [174].

The frequency of the CTCs differ by patients, but it is known that the frequency of the CTCs is often less than one per 1 ml of the peripheral blood. To effectively isolate the one CTC from the peripheral blood, there should be a minimized loss of the CTCs from the blood sample, which inhibits the samples from the FACS isolation. Therefore, if the CTCs could be enriched and analyzed thoroughly via image, it would be able to isolate the CTCs and proceed the post-processing for identifying the genomic aberration of the primary tumors.

CTCs can be enriched using the mechanical isolation methods such as microfluidic devices to increase the ratio of CTC to WBC from the blood plasma. Before applying the schematic to the human patient samples, I used mixed cell lines of HL-60 (promyeoloblasts) and MCF-7 (breast cancer) to mimic the human samples. Then, by using the anti-CD45 (FITC) and anti-pan cytokeratin (Cy5) to stain the HL-60 and MCF-7, respectively.

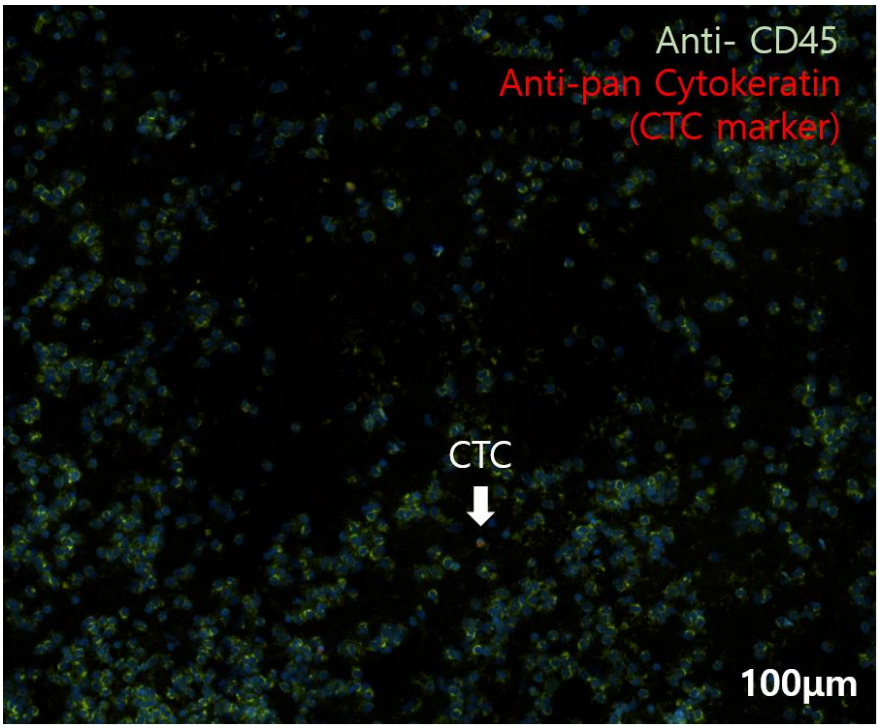


Figure 3.11 The image of the CTC samples among the peripheral blood.

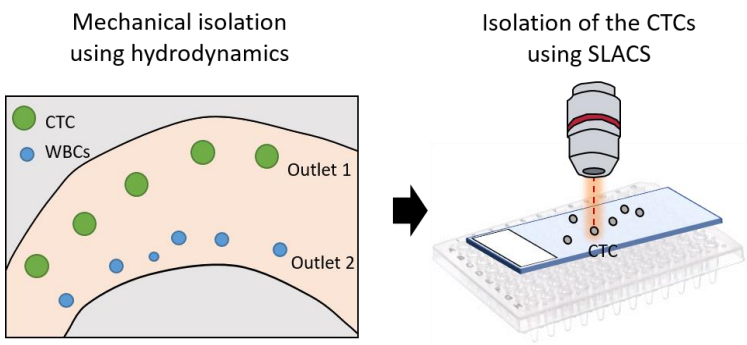


Figure 3.12 Isolation of CTCs from enriched CTCs

3.2.2. Development of the cell sorting algorithm

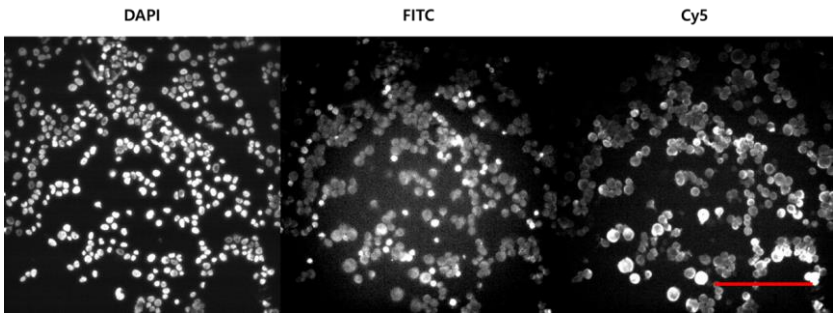


Figure 3.13 Image of each channels of mixed cell lines of HL-60 and MCF-7

Figure 2.14 shows the initial image of the fluorescence channels of DAPI, FITC, Cy5 which are DAPI, anti-CD45, anti-pan cytokeratin respectively. Using the cell profiler, which is an open source tool for analyzing the image of the cells, I first tried to measure the nucleus from the DAPI channels [184].

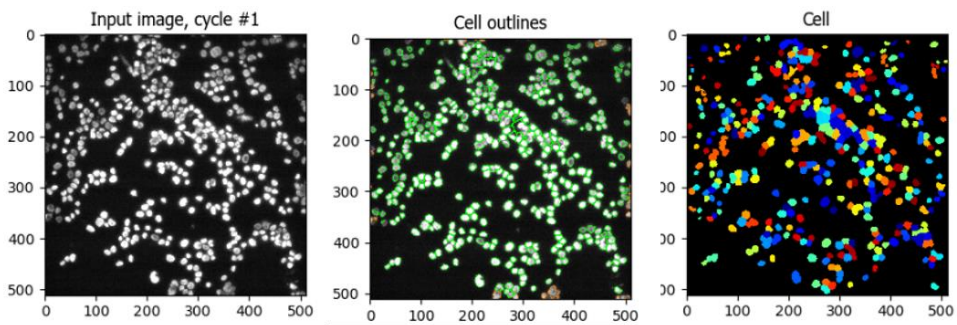


Figure 3.14 Nucleus detection from the DAPI image

After the validation of the nucleus detection, I then tried to extract and relate each fluorescent channels with the nuclei. To actively classify the HL-60 cell line

and MCF-7 cell line from the image, I developed the classification algorithm. First, I split the individual channels for each channels to specify the primary circular objects in each channel. Then I select the cytoplasm that match the nuclei to eliminate the non-specific bindings. The overall flow of the classification of mixed cell lines is described in Figure 2.16.

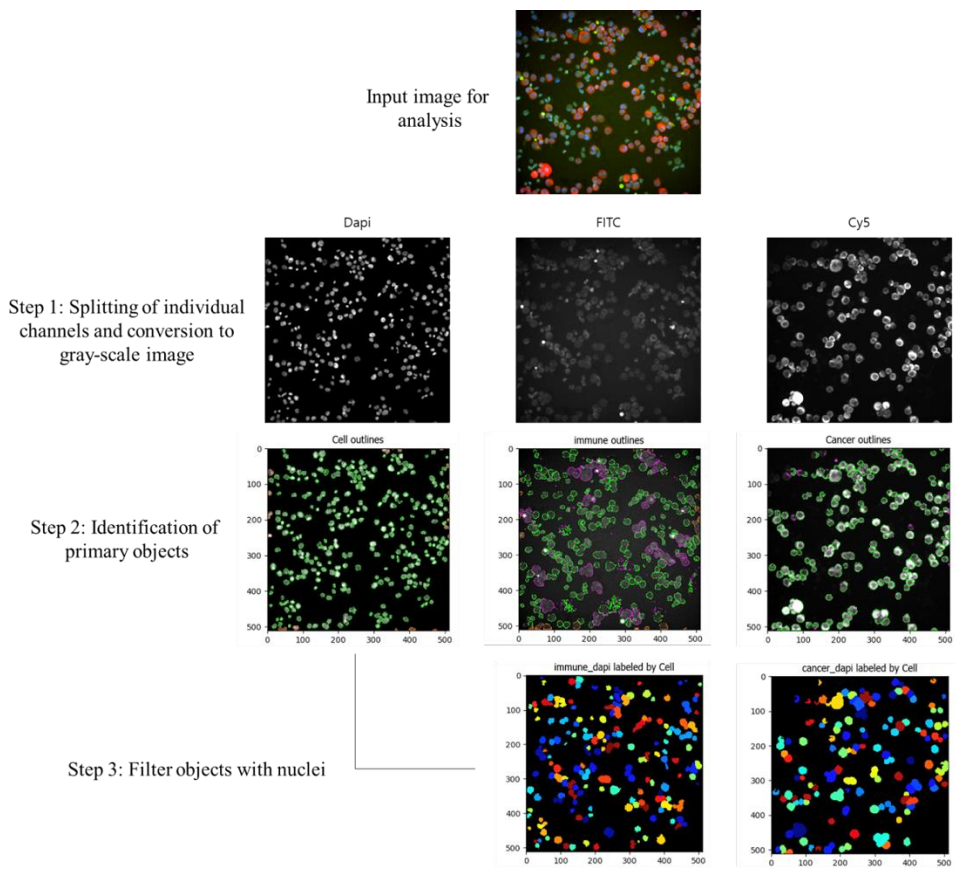


Figure 3.15 Classification algorithm and the resulting image of the classified cells.

Using the algorithm, it was able to identify the objects from the images,

however there were intrinsic problem regarding the fluorescence image. Unlike the Cy5 channel which corresponds to the cancer cell line, FITC channel showed not only the fluorophore that are attached to the anti-CD45, but also with the cancer cell lines. It is known that FITC channel which has the emission wavelength of 530nm is known to overlap with the natural fluorescence emitted by cells.

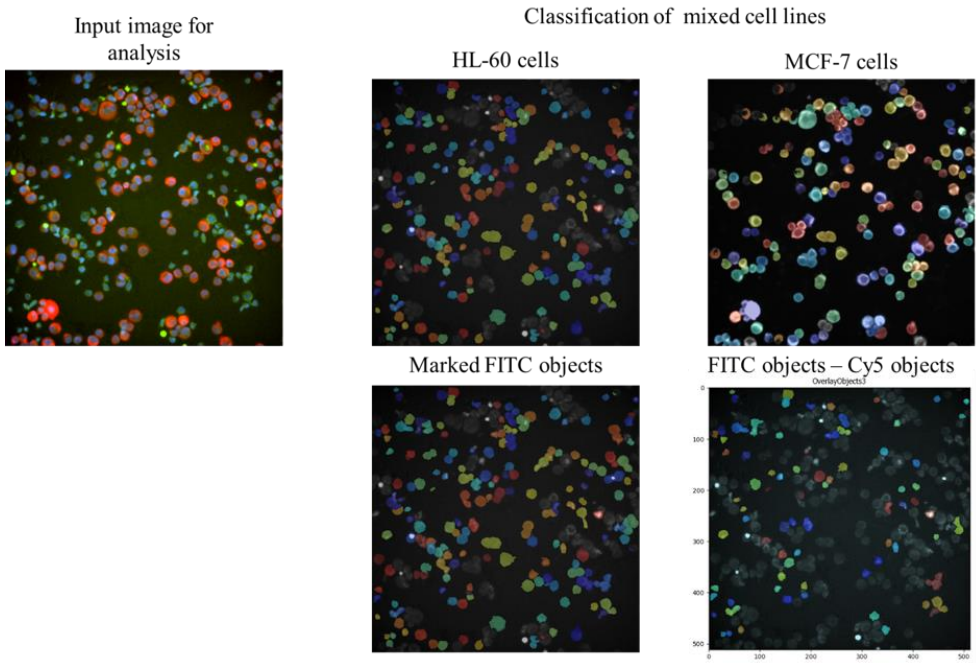


Figure 3.16 Optimization of filtering the autofluorescence from FITC channels

I then optimized the FITC channel analysis by filtering the natural fluorescence from the cancer cell lines which might be classified as the HL-60 cell lines. First, each cells were related with the parent cells which are detected by the DAPI channels. Then, by filtering the cells which shares the parent cells with the FITC cell lines, it

was able to preclude the cells which are initially marked by the Cy5 channel.

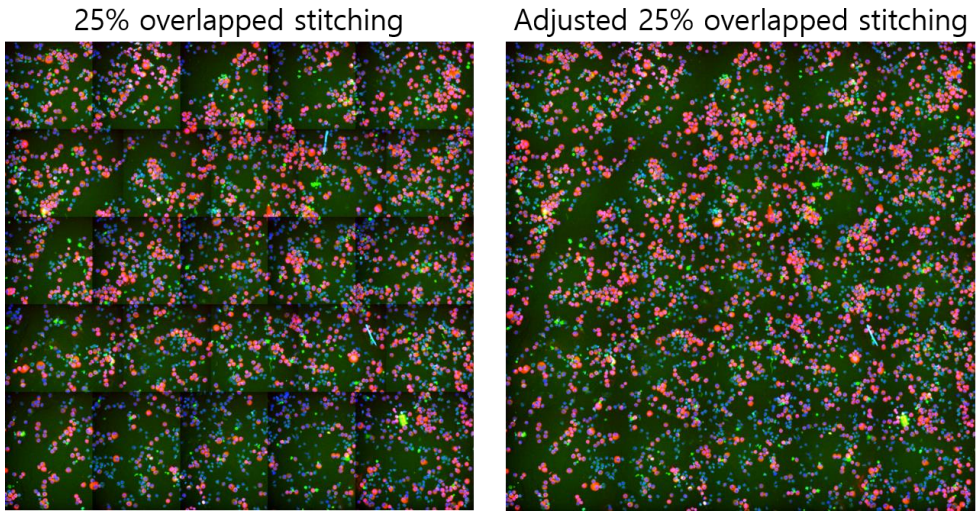


Figure 3.17 Adjusted stitching for large image scanning

Then, to increase the scale of the sorting, I imaged each fields and stitched each images to gain the image of the large area with high resolution. Each images were taken by 512 x 512 pixel image using the Hamamatsu CCD. Each pixels refer 1.62 μm in actual scale. The image condition had the non-uniform illumination of the light source, so stitched images were adjusted by extracting the maximum fluorescent image of the overlapped areas. However, it was difficult to filter out all the cancer cells from the FITC image in the large scale, due to the non-uniform imaging conditions, so I decided to plot the fluorescence profile of each objects and map the profiles to group each types of cells.

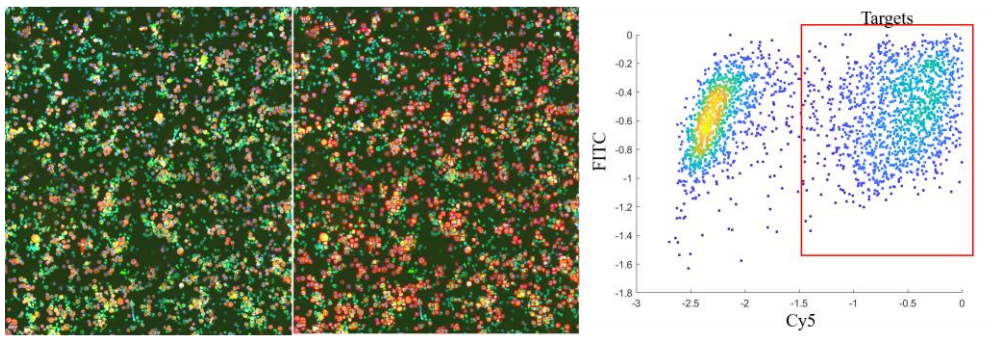


Figure 3.18 Selected areas of the isolation and corresponding fluorescence profiles of the cells

Red boxed cells are selected targets that are selected due to its fluorescence profile. It was difficult to directly extract the profiles from each cells, so by plotting the fluorescence profile of the each primary objects, it was able to differentiate each cells into two groups. Therefore, rather than depending on the certain fluorescent image, it was able to precisely select the target profiles by introducing the profile plots which are similar to the gating method of the FACS.

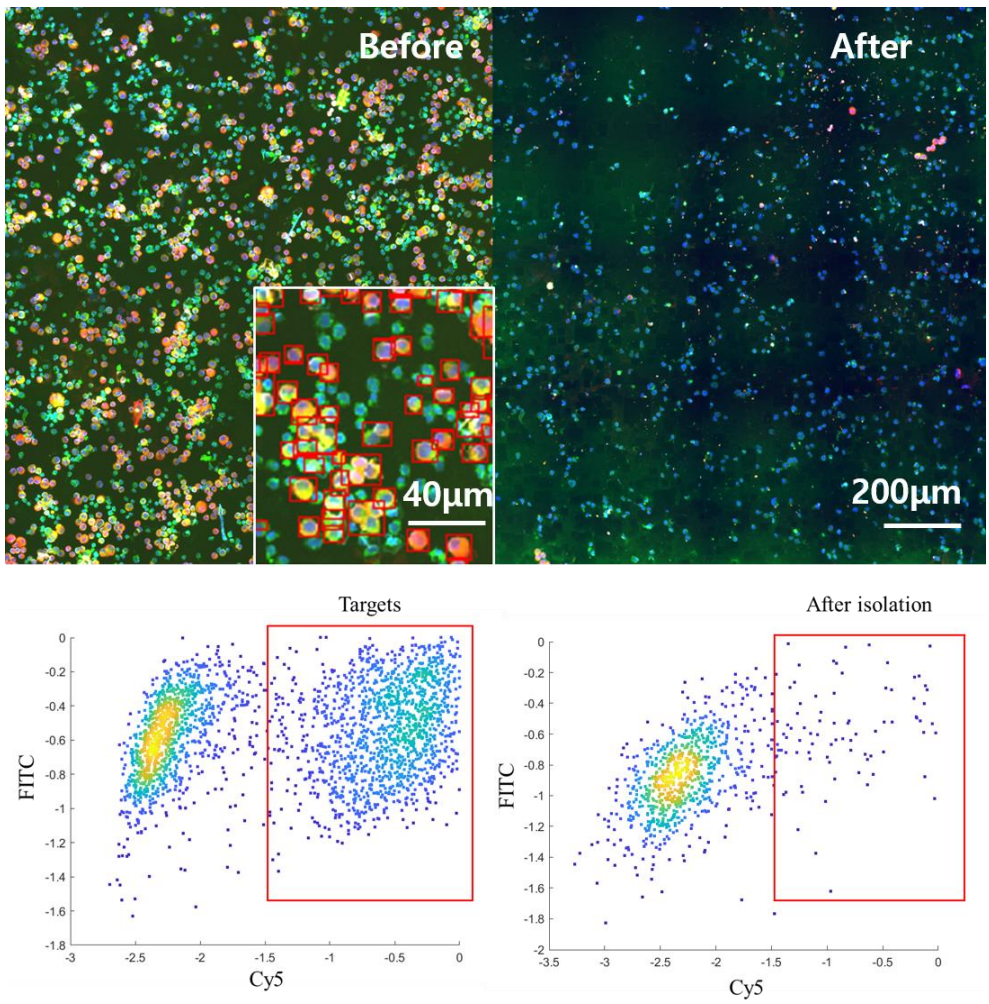


Figure 3.19 Isolation of the target cells that are pre-marked by the targeting software

It was possible to target the cancer cells that have higher intensity of the Cy5 channels. By pooling out the coordinates of the targets regions of the marked regions, it was possible to isolate the target cells using the SLACS device. It took about an hour to isolate 2500 cells, which was almost a target-per-second. When compared to

FACS, it had lower isolation speed, but instead it was possible to keep the spatial information within the regions.

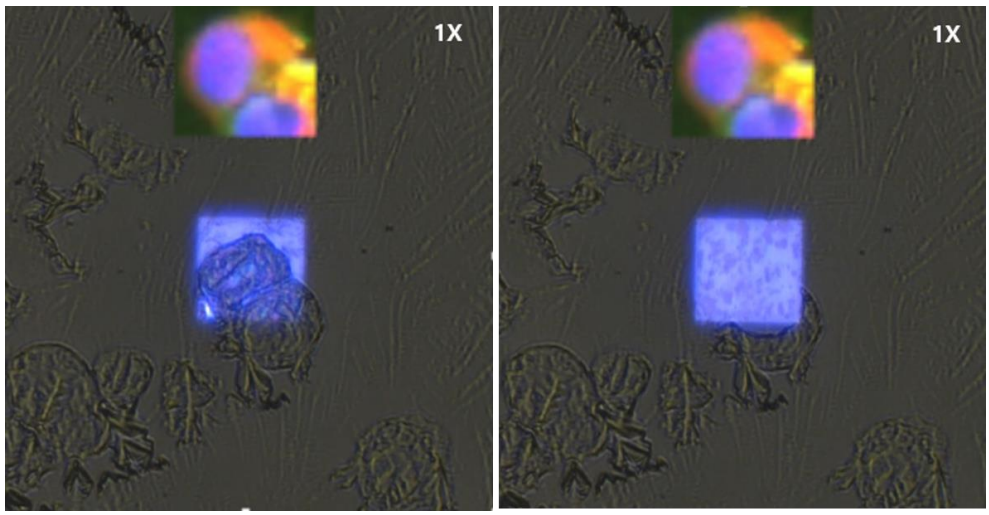


Figure 3.20 Real-time image of the isolation of target cells

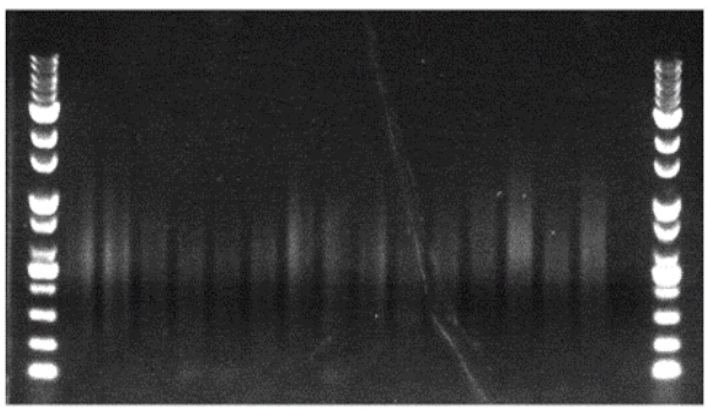
The image of the isolation target and the corresponding isolation process is shown in Figure 2.21. The target process could even control the size of the laser illumination spots, so that is could actually target the desired regions. The illumination of the UV laser was nano-second pulse and the most of time spent for the isolation was the time for moving the motorization stage.

3.2.3. The quality of the biomolecules from the isolated cells

To validate the performance of the cells isolated by the SLACS device, I first compared the DNA quality after isolating 1 cell, 10 cell and 100 cells, I ran the direct library preparation method for effectively amplifying the genomic molecules from

the cells. The data for quality control of the DNA molecules is visualized in Figure 3.22. Because the DNA molecules are amplified, there was no significant difference between the number of cells. It is verified that minimum of 10 ng/μl of the genomic molecules are obtained after the direct library preparation methods, which were fair enough for the next generation sequencing.

- **Data**



	Index
1: 1 cell	I711-I502
2: 1 cell	I711-I503
3: 1 cell	I711-I504
4: 1 cell	I711-I505
5: 1 cell	I711-I506
6: 10 cell	I711-I508
7: 10 cell	I711-I509
8: 10 cell	I711-I510
9: 10 cell	I712-I502
10: 10 cell	I712-I503
11: 100 cell	I712-I504
12: 100 cell	I712-I505
13: 100 cell	I712-I506
14: 100 cell	I712-I508
15: 100 cell	I712-I509
16:N	I712-I510

(ng/ul)

1	1	1	1	1	10	10	10	10	10	100	100	100	100	100	N
16.5	21.2	10.2	13.8	13.5	15.6	18.4	11.2	13.3	7.61	15.3	11.4	24.7	12.4	19.5	0.816

Figure 3.21 Quality control data of the isolated cells

The application for the full-length RNA sequencing was also tested to validate the preservation of the transcriptome features after the isolation of the cells. Using the Smart-seq2 [185] which follows poly T mediated reverse transcription of the mRNA to synthesize the cDNA, template switching, PCR amplification of the DNA/RNA hybrids, TN5 tagmentation of the NGS barcodes and PCR steps to

amplify the NGS libraries. It is confirmed that after the Smart-seq2 process, at least 10 ng/μl of the molecules were obtained in the most of the samples, and was confirmed through the gel electrophoresis data that genomic molecules had the constant band of the length.

• Qubit

(ng/ul)

1	1	1	1	1	10	10	10	10	10	100	100	100	100	N	N
4.61	42.4	34.2	19.8	25.3	23.9	16.2	18.2	28.6	25.9	16	6.69	15.9	2.53	2.17	1.62

• Gel electrophoresis

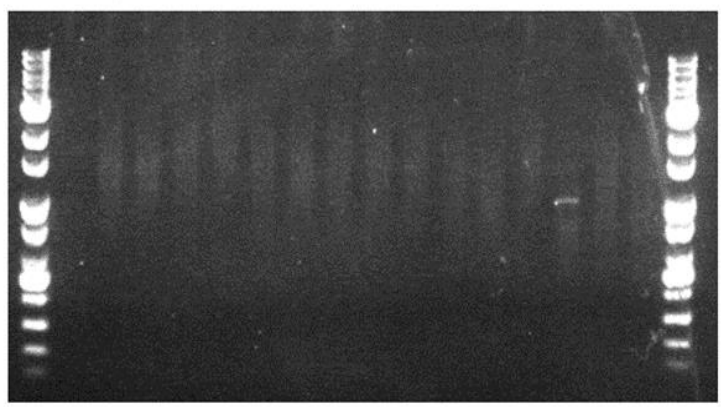


Figure 3.22 Quality control of the RNA after cell isolation using the SLA CS

3.2.4. Application to isolating the circulating tumor cells (CTCs)

Then I applied the image profiling process to the purified CTC image from the patient samples. To fully extract the CTCs from the blood samples, WBCs with the similar size were filtered together with the CTCs. Among them, purified solutions from the blood plasma were spread over the ITO glass using the Cytospin device,

which uses the centrifugal force to deposit the cells on the glass slide.

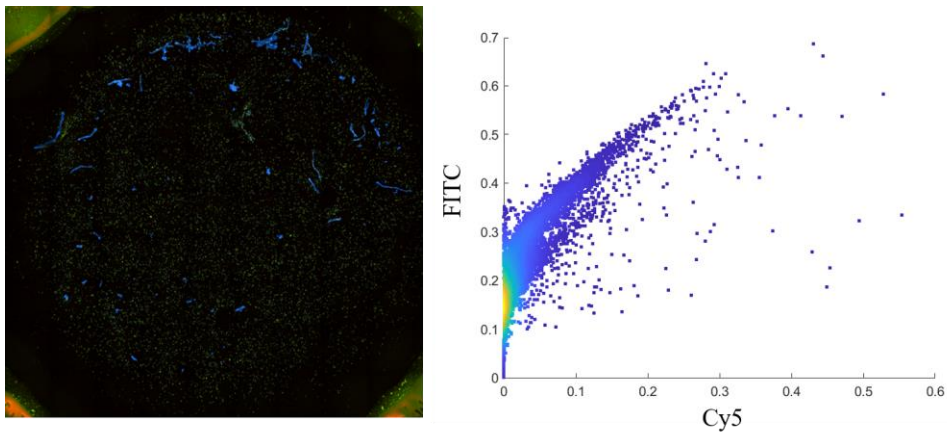


Figure 3.23 Image profile of the CTC sample

Unlike the cell line experiment, where cancer cells and promyeoloblasts were mixed in 1:1 concentrations, actual patient samples had lack of number of the CTCs among the samples. So by manual counting of the CTC cells, it took about 3-4 hours to profile 10,000 cells deposited on the cells. However, by using the program, it was able to profile the intensity of each cells, and further were able to select the targets using the gating informations. The CTC like features, where the cells have higher intensity of anti-pan cytokeratin profiles, were selected in the image, and the corresponding target image and target coordinates are extracted from the desired images.

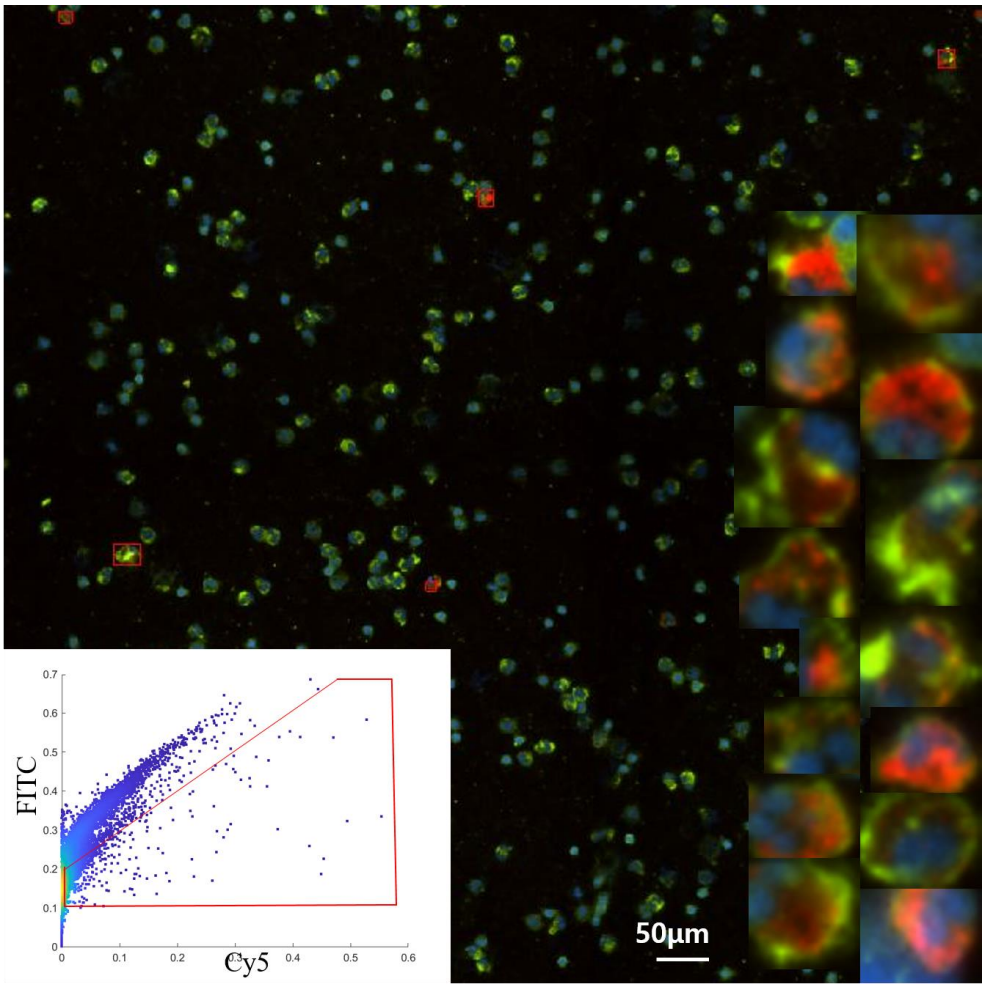


Figure 3.24 The image of the marked CTCs in the patient samples and selected targets are attached in the right side

3.3. Automated cell sorting for targeting the clinical tissue samples

3.3.1. Need for sorting the tumor regions for specific markers

It is known that tumor regions are consisted of various tumor subtypes. Therefore, increasing number of researches are proceeded in tumor regions to effectively analyze the tumor section in the spatial context to effectively inspect the tumor heterogeneity, cell-to-cell interactions, tumor microenvironment and to discover the novel therapeutic or diagnostic makers for the clinical actions [186]. Based on the spatial barcoding approaches or the *in situ* fluorescent imaging methods, the heterogeneity and tumor microenvironments are actively conducted, but biomarker discoveries are proceeded only in terms of the relative expression level of the transcripts due to the technical huddle.

To identify the novel tumor biomarkers, it is especially important to effectively identify the full length transcriptomes for in-depth characterization of the corresponding tumor regions. Spatially barcoded transcripts only targets the small fraction of the whole transcripts due to the limitation in NGS cost. On the other hand, if it is possible to extract out the regions with target biomarkers and proceed the full-length sequencing in the corresponding regions, it would be possible to effectively identify the transcripts with less sequencing cost.

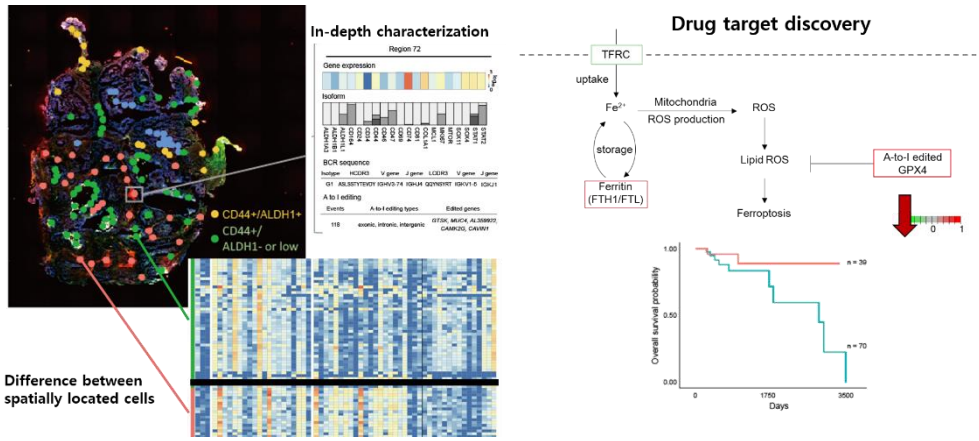


Figure 3.25 Therapeutic target discovery using SLACS (Reprinted from [173])

The study conducted using SLACS have shown is possibility of identifying the novel therapeutic markers by connecting the immunostaining data with its full-length transcriptome information. In the study, two cancer stem cell like markers, CD44 and ALDH1, was stained on the tissue sample of the patient sample of residual tumor who have been treated with neoadjuvant chemotherapy. Cancer stem cells means that it still has the stem-cell like behavior which develops the tumor even after the therapy. Therefore, the study was conducted to extract the therapeutic markers which are correlated with the residual tumor. Separated with 4 different groups, it was possible to identify the new target of A-to-I edited GPX4 which showed higher expression and A-to-I edited frequencies in the tumor stem cell like regions. When compared to the TCGA data, it was confirmed that patient with A-to-I edited GPX4 showed the worse clinical outcome. In the suggested study, it was able to extract out the novel tumor markers, but had the limitation in that the decision of the interpreting the

immunostaining data was conducted without the certain standard, and the isolation was proceeded manually.

3.3.2. Cell sorting of the target markers in clinical samples

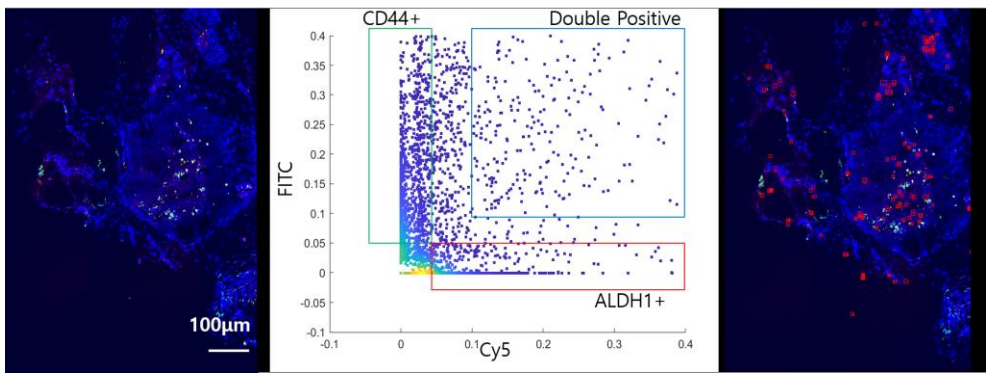


Figure 3.26 The result of target sorting using the automatic target selection program

I stained the tumor tissue using immunostaining with the markers of CD44 and ALDH1. Image analysis using Cellprofiler allowed us to profile the intensity of each fluorescence channel in each cell. By mapping the corresponding profiling data, similar graphs such as FACS gating plot can be obtained. Then, from the data, I extracted the coordinates of four different groups of targets, CD44 positive, ALDH1 positive, double positive and double negative. Then, I ran the laser-activated transfer of the target cells to the 8-strip of 96 well tubes.

It is known that the single cell contains 30 pg to 100 pg of transcriptomic contents.Unlike cell lines where the whole cells are deposited on the ITO glass, cells

in the tissue slides are sectioned during the cryosection process. Therefore, it is important to amplify the mRNA before the sequencing, and if the content is insufficient for the sequencing, same typed cells can be gathered to be sequenced. I first compared the genomic quality of the 5, 20, and 100 cells of the tissue are collected in a tube and ran Smart-seq2 to obtain the full-length transcriptome from the tissue samples [185].

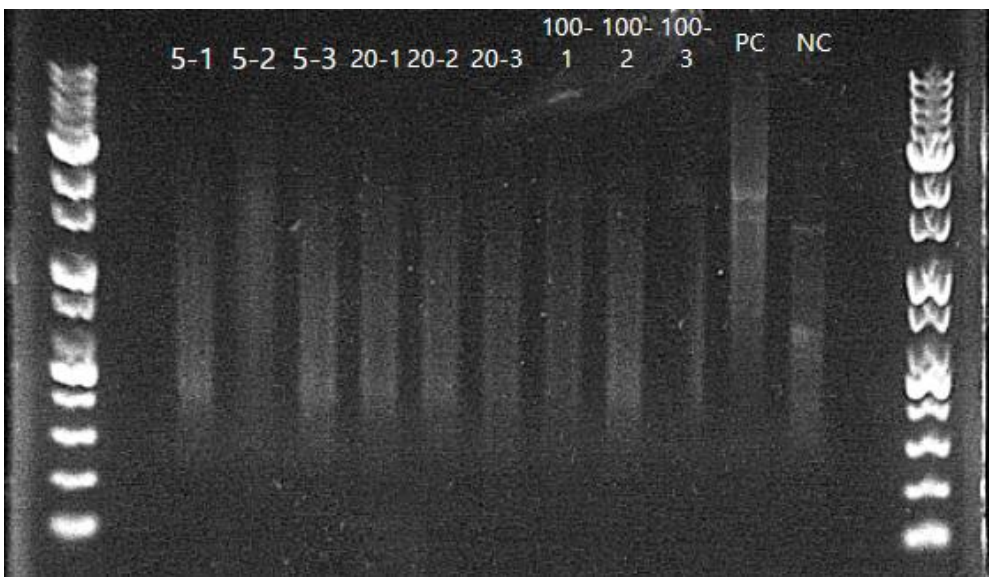


Figure 3.27 The RNA quality of the isolated cells (5, 20, 100 cells triplicate each)

Chapter 4.

Integrated spatial profiling technologies

In this chapter, I will describe the precise selection of the region-of-interest after spatially profiled biological samples. While SLACS is more focused on in-depth characterization, there are spatial landscape technologies that has advantage in profiling the expressed genes in large area. After profiling of the spatial expression of target genes or most-expressed genes, it is possible to apply the in-depth characterization analysis or to add other modality such as multi-omics approach.

4.1. Integration with the other spatial landscape profiling technologies

4.1.1. *In situ* sequencing for profiling the spatial landscape

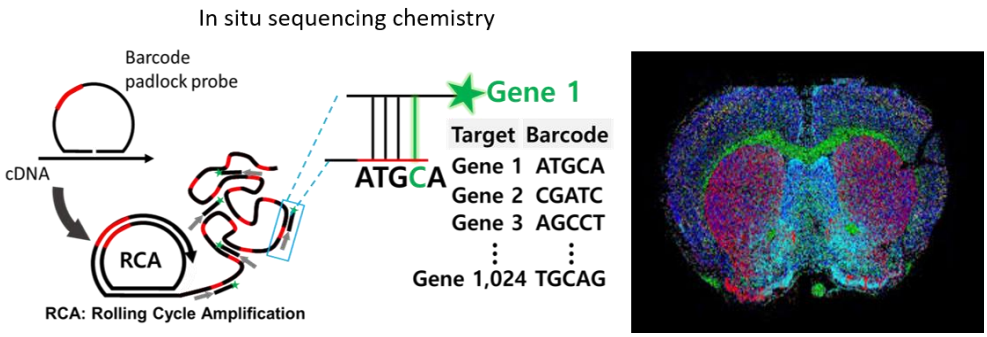


Figure 4.1 Schematic of the *in situ* sequencing chemistry for mapping the spatially expressed transcriptomes [53]

In situ sequencing is one of the state-of-art technology to profile the spatially expressed transcripts using the fluorescently labelled probes to sequence the barcodes of the padlock probes. It implements the padlock probes to amplify the genetic molecules for the sufficient amplification of the fluorescence signals to be detected. The first *in situ* sequencing method was published in 2013 and are now commercialized by 10X genomics as Xenium. In the primary chemistry, it could either quantify the target genes or to detect the single-nucleotide variants using gap filling strategy. The chemistry is followed by (1) Fixation of the mRNA, (2) reverse transcription, (3) padlock hybridization to the cDNA, (4) Rolling circle amplification (RCA) of the padlock probes (5) sequencing of the padlock probes by fluorescence

probes.

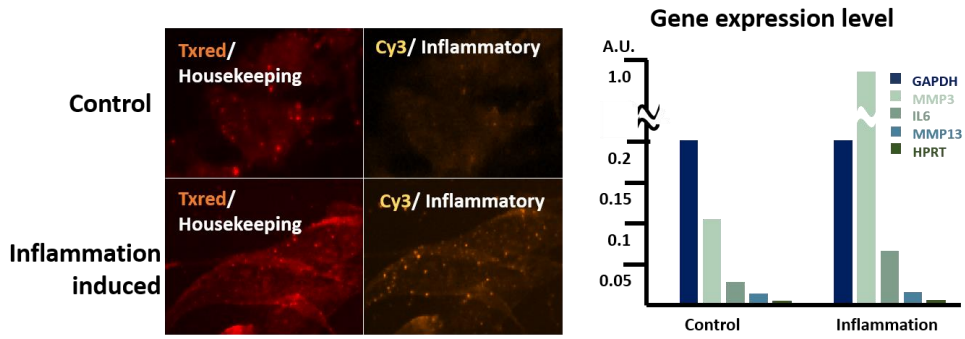


Figure 4.2 Validation of *in situ* sequencing in control and inflammation induced sample

Initially, I tested whether the *in situ* sequencing could differentiate the two different type of samples in the mouse articular chondrocyte. Inflammation was induced by the interleukin-1 beta treatment. The product of the *in situ* sequencing is the group of fluorescence blobs which are similar to the product of the FISH/ As seen in the Cy3 channel which contains the inflammatory genes, the number of blobs increase in the inflamed conditions. Padlock probes for Gapdh, Mmp3, Il6, Mmp13, and Hprt was designed for targeting each genes. Gapdh, Hprt was used as a housekeeping gene for normalizing the transcripts and Mmp3, Mmp13, Il6 are genes related to inflammatory functions which are known to increase in inflamed condition.

To enhance the accuracy and specificity of the gene annotations, the image quality needed the development by enhancing the signal-to-noise ratio and the number of the RCPs.

Initially, the rolling circle amplified products (RCPs) showed low signal-to-noise ratio, resulting in the missing outs of the image analysis. Therefore, to enhance the quality of the assay results, I made some optimization to the protocol of the *in situ* sequencing as follows. (1) Reaction time increment: to effectively increase the efficiency of the fluorescence probe and padlock hybridization steps, (2) Fixation protocol optimization: initial protocol of the fixation method utilized 4% Paraformaldehyde (PFA) to fix the transcripts to the cells, but changed the protocol to 4% formaldehyde for stronger fixation of the expressed transcripts. (3) Reaction chamber optimization: the initial protocol from the SciLife Lab implements the chamber that can stick to the desired regions, however the adhesive that chamber utilized results in the aggregation of the reaction buffers, therefore changed the reaction chamber by designing the wall to confine the raction buffers made up of Polydimethylsiloxane (PDMS). By doing so, the number of detected blobs and the signal-to-noise ratio showed the enhancement.

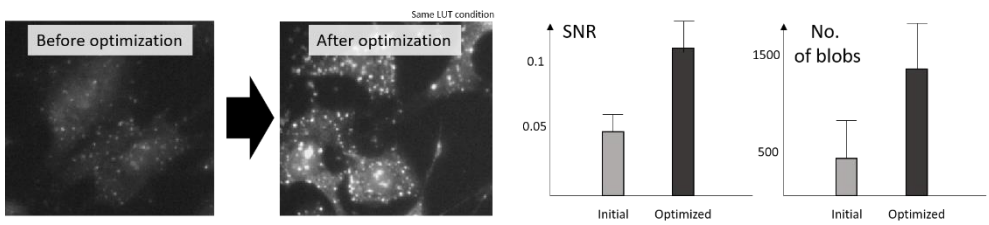


Figure 4.3 Enhancement of the *in situ* sequencing assay quality

4.1.2. Gene annotation and analysis of *In situ* sequencing data

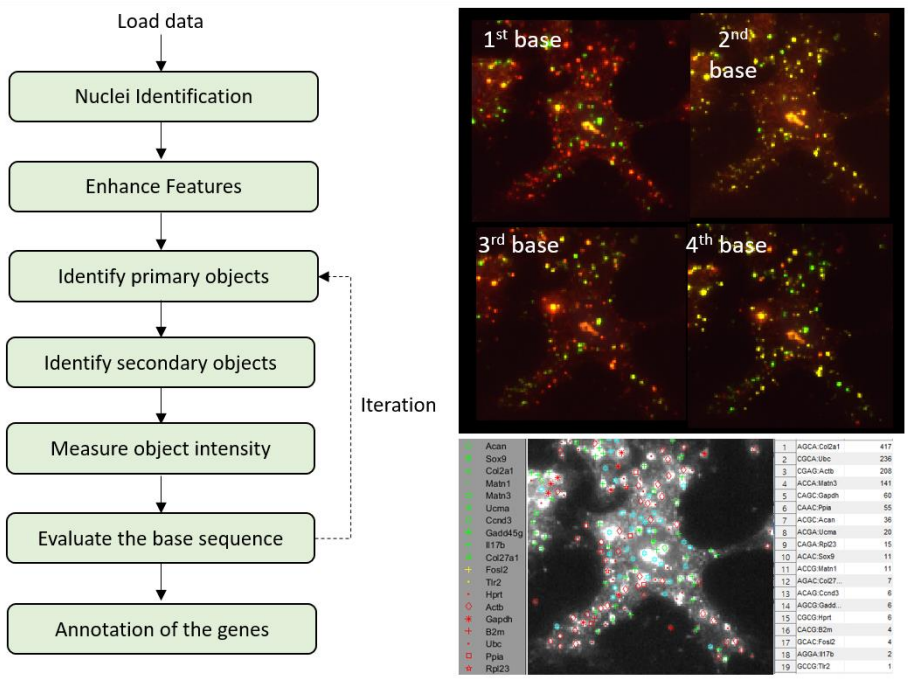


Figure 4.4 Gene annotation program from the result of the *in situ* sequencing assay

To effectively annotate the transcripts using the *in situ* sequencing data, I developed the gene annotation program which can determine barcodes from the series of image data built on the program developed in Prof. Nilsson's lab. The pipeline of the gene annotation is listed in Fig 3.31. First of all, based on the series of the fluorescence channel image, the images are proceeded through (1) nuclei identification, (2) enhance features: Fluorescence image contains the noise due to the cell background, non-specific binding, and imaging conditions. Blobs are shown as the group of high intensity pixels, so it is important to enhance the signal quality

by Gaussian smoothing, (3) identification of primary objects: based on the smoothed image, all of the RCPs are detected, (4) identification of secondary objects: based on the RCP images, the specific blobs are classified to the matching fluorescence channels, (5) measure object intensity: the intensity of the each fluorescent channels are profiled, (6) evaluate the base sequence: the base sequence of each blobs are determined according to the maximum intensity of fluorescence channels, (7) iteration of the barcode determination steps: *in situ* sequencing implements the series of fluorescence probe hybridization steps to read the barcode sequence of each blobs, therefore is critical to align the each sequence to identify the barcode sequence, (8) barcode annotation: based on the determined sequence for each iteration steps, the barcode sequence can be annotated as well as the corresponding genes.

The image of the *in situ* sequencing result of the each rounds are shown in Fig 3.32. As each iteration step proceeds, each blobs change the belonging fluorescence channels and by tracking them, it is possible to determine its barcoding sequence and corresponding gene annotations. The result of the *in situ* sequencing is can be annotated on the image of the *in situ* assay results.

4.2. Multi-omics analysis in integrated spatial profiling

4.2.1. *In situ* sequencing in two different groups

According to the analyzed data, it was able to plot the heatmap of the expressed transcripts of the two different groups (control vs inflammation induced group) and verified that there was a critical difference in A (Anabolic genes), C (Catabolic genes)

compared to the house keeping genes. By drawing the principle component analysis (PCA) plots, it was able to differentiate that two groups showed the different features using the *in situ* sequencing analysis. Targeting the specific regions of the interest using the in situ sequencing are also possible.

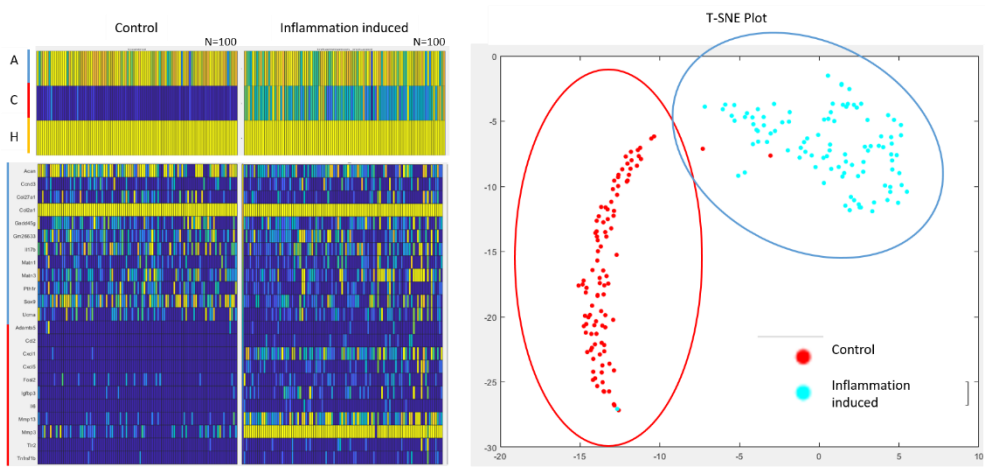


Figure 4.5 Analysis of the gene expression difference between control and inflammation induced samples

In situ sequencing has an advantage that it can specify the cells of interest using the transcriptome information. For instance, it is capable of mapping the overall expressed transcripts directly on the clinical tissues, and by using the *in situ* sequencing information, it is available to isolate the specific regions of interest which are important in the pathological context.

4.2.2. Multi-omics profiling by integrating in situ sequencing to its genomic aberrations

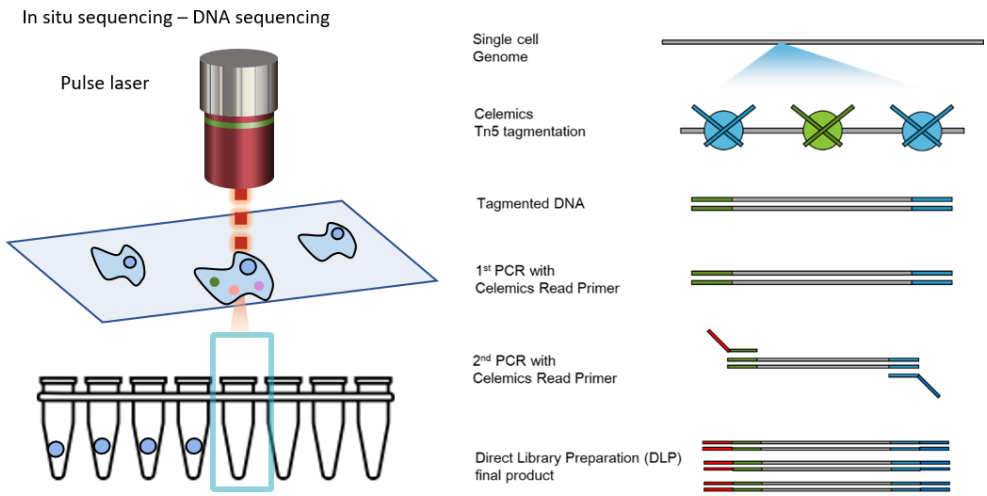


Figure 4.6 Direct library preparation of the isolated cells after in situ sequencing

It is possible to target the desired cells of which transcriptome expressions are profiled using in situ sequencing. Then, it was able to proceed the direct library preparation for amplifying the DNA strands for whole genome sequencing. By proceeding the DNA sequencing after spatial landscape transcriptome profiling technologies, it is possible to connect the genomic aberrations with its transcriptome profiles.

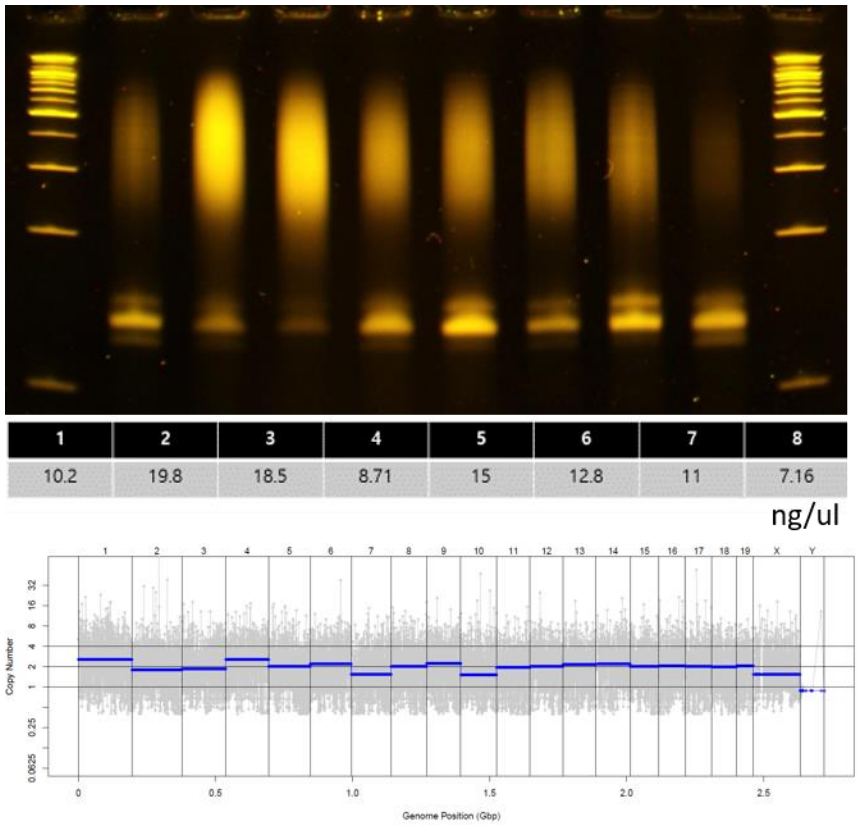


Figure 4.7 Result of the copy number variation of the in situ sequencing sample

The quality of the DNA strand in the nucleus might not be sufficient for the assay, resulting in the dimer sequences of 150 base pair length. Therefore, I proceeded the gel size selection excluding the 150bp length sequences and proceeded the direct library preparation DNA sequencing The reason of the resulted dimers might be due to the double stranded DNA of the sequencing probes. When reading the barcode sequence using the fluorescence probe, it constructs the double strand DNA, and that might have been amplified during the genomic molecule

amplification process. It still needs the enhancement of the chemistries, but still it was able to proceed the copy number variation analysis in the transcriptome mapped samples. Furthermore, it would be able to analyze the efficacy of the drug by matching the efficacy of drugs to the genomic aberrations.

4.2.3. Combination with other spatial profiling technologies

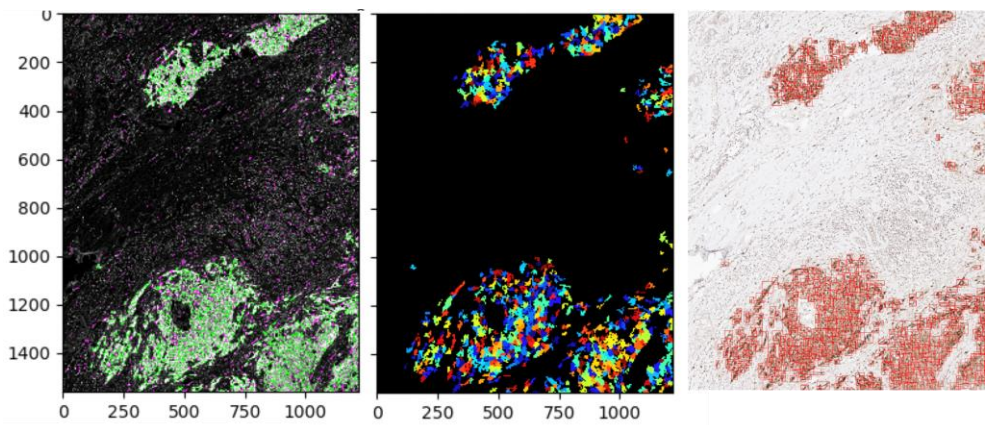


Figure 4.8 ROI selection on immunohistochemistry (IHC)-stained sample

SLACS utilizes the image to guide the ROI selection from the tissue, therefore, is possible to apply any types of staining method where the regions can be specified by staining. There is no need for selecting the compatible antibody or to amplify the signal for analyzing the ROIs. Other staining methods or spatial transcriptomics profiling methods such as multiplex FISH methods or Visium can assist guiding the target cells from the tissue and further be connected with the high quality of genetic assays.

Chapter 5.

Conclusion and Discussion

In this chapter, the proposed platform is summarized. The need and importance of the developed device are suggested and, the potential of the platform and its future work is described.

5.1.1. Summary of the dissertation

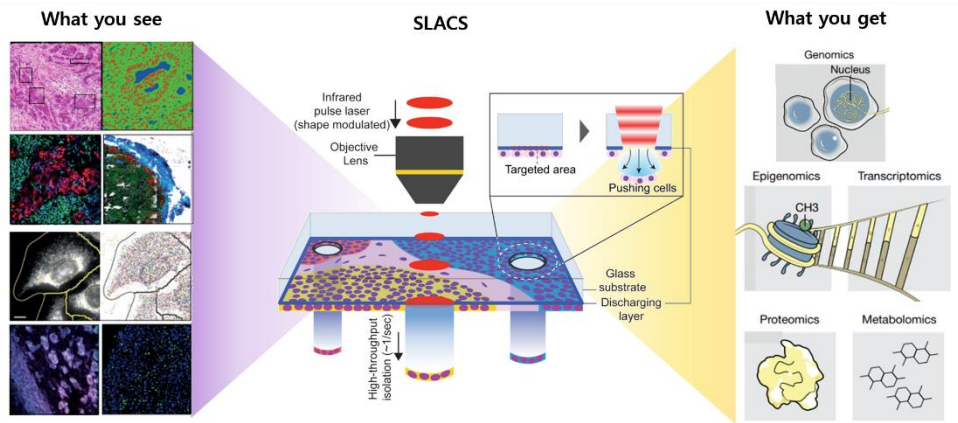


Figure 5.1 Suggested SLACS platform for automated cell sorting which can link the genetic molecules with its spatial context

In this dissertation, I described the importance and need for spatial assays linked with the biomolecular assays. As the pathology involves complex mechanism within various types of cells, it is important to target the specific region of interest and proceed the in-depth characterization to understand the cellular mechanisms and discover the novel biomarkers regarding the pathology. Previously developed SLACS device is capable of isolating the regions of interests, but involves the manual determination and targeting from the molecular-marked images.

I validated the automated detection of the targets using the encoded microparticles, and linked the target coordinated to the SLACS device which illuminates the target coordinates to isolate the corresponding regions. Then, target cell type retrieval from mixed cell lines was validated, further expanded to the

application of CTC retrieval. Also, the tissue sample which are stained with the CD44 and ALDH1 stem cell markers are analyzed and each groups were determines using the automated pipeline to retrieve each groups SLACS was applied to other spatial transcriptomics landscaping technologies such as *in situ* sequencing to determine the cells in high definition and linked transcriptome features to the genomic aberration such as copy number variations.

By integrating the image processing and automatic detection and isolation of targets, it is able to isolate up to 2500 targets in an hour. In this dissertation, I have suggested the concept of the automation and validated the isolation and the quality of the genomic molecules from the isolated samples. By highly increasing the throughput of the platform, it will broaden the applications and will provide the tool for the spatial omics assay.

5.1.2. Limitation of the technology

In current state of the technology, it is only applicable in the fluorescent-labelled image, and are willing to expand the automated targeting to the H&E stained tissue samples which are used in the clinical practice Unlike the fluorescence labelling methods which has an obvious criteria for defining the targets, H&E images contains complex criteria which might even differ within the pathologists. Increasing researches are introducing the deep learning to train and extract the pathologically relevant regions. Therefore, I am trying to expand the platform to be

5.1.3. The impact of the suggesting technology

As the cost of next-generation sequencing gets lower, discovering the landscape of the genome, transcriptome, epigenome and proteome and exploring regions of interest has been made possible by thriving spatial technologies. Since the Human Genome Project began in 1990, the Cancer Genome Anatomy Project (1997), Cancer Genome Atlas (2006), Human Cell Atlas (2016) and many other projects have attempted to build a database or atlas landscape of the cell composition of the human cancer. The next step is to construct a spatial or pathological atlas that comprehensively maps the cellular landscape in the highly heterogeneous tumor microenvironment using the genomic features. However, whether it is a large-scale spatial omics profiling technologies or a region-of-interest based spatial profiling technologies, the applicability of both technologies is not limited to just atlasing the cellular compositions.

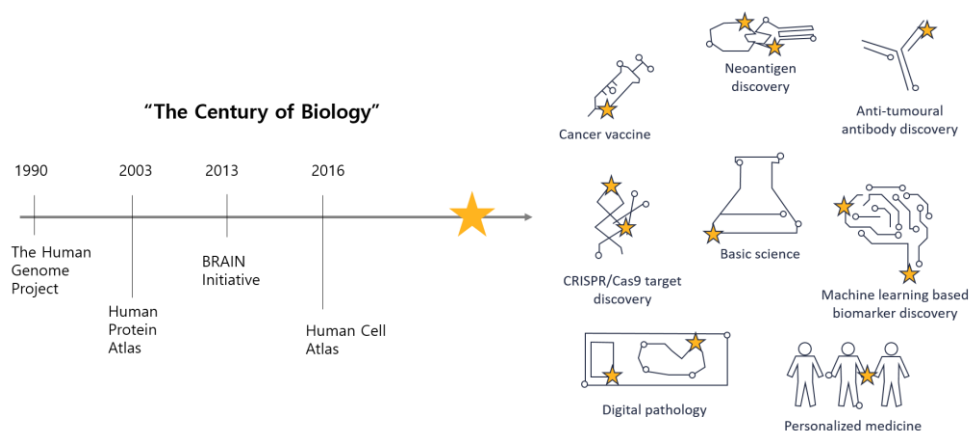


Figure 5.2 The next generation therapeutic approach using the spatial omics profiling

technologies

I believe that the spatial omics information in pathology will lead the next generation diagnostics and therapeutics. Increasing number of researches have already begun implementing the spatial omics combined with machine learning to discover diagnostic biomarkers and apply it to digital pathology [187]. Furthermore, the discovery of transcribed sequences found in cancer can be used in therapeutic target discovery using mRNA-based cancer vaccines, neoantigen-targeted chimeric antigen receptor T-cell (CAR-T-cell) therapy, gene editing therapies including CRISPR/Cas9, RNA interference, and many other treatments. In addition, immune cell profiling of the tumor infiltrating cancer cells present in the tumor regions poses the great potential in the discovery of anti-tumor antibodies. The potential of spatial omics continues to expand and in the near future, these next generation spatial omics tools will thrive the therapeutics and diagnostics developments.

Bibliography

[1] G. B. Faguet, “A brief history of cancer: Age-old milestones underlying our current knowledge database,” *Int. J. Cancer*, vol. 136, no. 9, pp. 2022–2036, May 2015.

[2] R. Hooke, *Micrographia: Or Some Physiological Descriptions of Minute Bodies Made by Magnifying Glasses, with Observations and Inquiries Thereupon*. Courier, 2003.

[3] M. J, “Uber den feineran Bau and die For-man der Krankhauten Geschwulste,” *Berlin G Reimer*, 1838.

[4] S. I. Hajdu, “The First Tumor Pathologist,” 2004.

[5] S. I. Hajdu, “A note from history: landmarks in history of cancer, part 3,” *Cancer*, vol. 118, no. 4, pp. 1155–1168, Feb. 2012.

[6] L. A. C. King, D Friday .; King, “A Brief Historical Note on Staining by Hematoxylin and Eosin,” *Am. J. Dermatopathol.*, vol. 8, no. 2, p. 168, 1986.

[7] H. C. Cook, “Tinctorial methods in histology,” *J Clin Pathol*, 1997.

[8] Y. Yuan, “Spatial Heterogeneity in the Tumor Microenvironment,” *Cold Spring Harb. Perspect. Med.*, vol. 6, no. 8, Aug. 2016.

[9] A. H. Coons, H. J. Creech, and R. N. Jones, “Immunological Properties of an Antibody Containing a Fluorescent Group.*:,” *https://doi.org/10.3181/00379727-47-13084P*, vol. 47, no. 2, pp. 200–202, Nov. 1941.

[10] G. L. Le Bouvier, “The Heterogeneity of Australia Antigen,” *Source J. Infect. Dis.*, vol. 123, no. 6, pp. 671–675, 1971.

[11] D. Huber, L. Voith von Voithenberg, and G. V. Kaigala, “Fluorescence in situ hybridization (FISH): History, limitations and what to expect from micro-scale FISH?,” *Micro Nano Eng.*, vol. 1, pp. 15–24, Nov. 2018.

[12] B. Liegl *et al.*, “Heterogeneity of kinase inhibitor resistance mechanisms in GIST,” *J. Pathol.*, vol. 216, no. 1, pp. 64–74, Sep. 2008.

[13] L. Hu *et al.*, "Fluorescence in situ hybridization (FISH): An increasingly demanded tool for biomarker research and personalized medicine," *Biomark. Res.*, vol. 2, no. 1, pp. 1–13, Feb. 2014.
[14] F. Sanger, S. Nicklen, and A. R. Coulson, "DNA sequencing with chain-terminating inhibitors.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 74, no. 12, pp. 5463–5467, 1977.
[15] E. S. Lander *et al.*, "Initial sequencing and analysis of the human genome," *Nat. 2001 4096822*, vol. 409, no. 6822, pp. 860–921, Feb. 2001.
[16] L. Hood and L. Rowen, "The human genome project: Big science transforms biology and medicine," *Genome Med.*, vol. 5, no. 9, pp. 1–8, Sep. 2013.
[17] "The Cancer Genome Atlas." [Online]. Available: https://www.cancer.gov/tcga. [Accessed: 01-Sep-2022].
[18] L. A. Herzenberg, R. G. Sweet, and L. A. Herzenberg, "Fluorescence-activated cell sorting," *Sci. Am.*, vol. 234, no. 3, pp. 108–117, 1976.
[19] E. Z. Macosko *et al.*, "Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets," *Cell*, vol. 161, no. 5, pp. 1202–1214, May 2015.
[20] F. Tang *et al.*, "mRNA-Seq whole-transcriptome analysis of a single cell," *Nat. Methods 2009 65*, vol. 6, no. 5, pp. 377–382, Apr. 2009.
[21] J. Almagro, H. A. Messal, A. Elosegui-Artola, J. van Rheenen, and A. Behrens, "Tissue architecture in tumor initiation and progression," *Trends in Cancer*, vol. 8, no. 6, pp. 494–505, Jun. 2022.
[22] Y. Wu *et al.*, "Spatiotemporal Immune Landscape of Colorectal Cancer Liver Metastasis at Single-Cell Level," *Cancer Discov.*, vol. 12, no. 1, pp. 134–153, Jan. 2022.
[23] A. Andersson *et al.*, "Spatial deconvolution of HER2-positive breast cancer delineates tumor-associated cell type interactions," *Nat. Commun. 2021 121*, vol. 12, no. 1, pp. 1–14, Oct. 2021.
[24] R. Elaldi *et al.*, "High Dimensional Imaging Mass Cytometry Panel to Visualize the Tumor Immune Microenvironment Contexture," *Front. Immunol.*, vol. 12, p. 1254, Apr. 2021.
[25] J. C. Wortman *et al.*, "Spatial distribution of B cells and lymphocyte clusters as a predictor of triple-negative breast cancer outcome," *npj Breast Cancer 2021 71*, vol. 7, no. 1, pp. 1–13, Jul. 2021.
[26] S. Han *et al.*, "Single-cell profiling of microenvironment components by spatial localization in pancreatic ductal adenocarcinoma," *Theranostics*, vol. 12, no. 11, p. 4980, 2022.
[27] R. Y. Ma, A. Black, and B. Z. Qian, "Macrophage diversity in cancer revisited in the era of single-cell omics," *Trends Immunol.*, vol. 43, no. 7, pp. 546–563, Jul. 2022.

[28] I. Dagogo-Jack and A. T. Shaw, “Tumour heterogeneity and resistance to cancer therapies,” *Nat. Rev. Clin. Oncol. 2017 152*, vol. 15, no. 2, pp. 81–94, Nov. 2017.
[29] A. Kulasinghe, J. Monkman, E. T. Shah, N. Matigian, M. N. Adams, and K. O’Byrne, “Spatial Profiling Identifies Prognostic Features of Response to Adjuvant Therapy in Triple Negative Breast Cancer (TNBC),” *Front. Oncol.*, vol. 11, p. 1, Jan. 2021.
[30] K. Schmelz *et al.*, “Spatial and temporal intratumour heterogeneity has potential consequences for single biopsy-based neuroblastoma treatment decisions,” *Nat. Commun. 2021 121*, vol. 12, no. 1, pp. 1–13, Nov. 2021.
[31] V. Marx, “Method of the Year: spatially resolved transcriptomics.”
[32] S. M. Lewis *et al.*, “Spatial omics and multiplexed imaging to explore cancer biology,” *Nat. Methods 2021 189*, vol. 18, no. 9, pp. 997–1012, Aug. 2021.
[33] Q. Zhu, S. Shah, R. Dries, L. Cai, and G. C. Yuan, “Identification of spatially associated subpopulations by combining scRNAseq and sequential fluorescence in situ hybridization data,” *Nat. Biotechnol. 2018 3612*, vol. 36, no. 12, pp. 1183–1190, Oct. 2018.
[34] M. Asp, J. Bergenstråhle, and J. Lundeberg, “Spatially Resolved Transcriptomes—Next Generation Tools for Tissue Exploration,” *BioEssays*, vol. 42, no. 10, pp. 1–16, 2020.
[35] Q. Li, X. Zhang, and R. Ke, “Spatial Transcriptomics for Tumor Heterogeneity Analysis,” *Front. Genet.*, vol. 0, p. 1573, Jul. 2022.
[36] R. H. Singer and D. C. Ward, “Actin gene expression visualized in chicken muscle tissue culture by using in situ hybridization with a biotinated nucleotide analog.,” *Proc. Natl. Acad. Sci.*, vol. 79, no. 23, pp. 7331–7335, Dec. 1982.
[37] A. M. Femino, F. S. Fay, K. Fogarty, and R. H. Singer, “Visualization of single RNA transcripts in situ,” *Science (80-.).*, vol. 280, no. 5363, pp. 585–590, Apr. 1998.
[38] J. G. Gall and M. L. Pardue, “Formation and detection of RNA-DNA hybrid molecules in cytological preparations.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 63, no. 2, pp. 378–383, Jun. 1969.
[39] J. M. Levsky, S. M. Shenoy, R. C. Pezo, and R. H. Singer, “Single-cell gene expression profiling,” *Science (80-.).*, vol. 297, no. 5582, pp. 836–840, Aug. 2002.
[40] R. W. Dirks *et al.*, “3’-End Fluorochromized and Haptenized Oligonucleotides as in Situ Hybridization Probes for Multiple, Simultaneous RNA Detection’,” 1991.
[41] A. Raj, P. van den Bogaard, S. A. Rifkin, A. van Oudenaarden, and S. Tyagi, “Imaging individual mRNA molecules using multiple singly

labeled probes," *Nat. Methods 2008 510*, vol. 5, no. 10, pp. 877–879, Sep. 2008.
[42] A. M. Femino, F. S. Fay, K. Fogarty, and R. H. Singer, "Visualization of single RNA transcripts in situ," *Science (80-.).*, vol. 280, no. 5363, pp. 585–590, Apr. 1998.
[43] F. Wang *et al.*, "RNAscope: A Novel in Situ RNA Analysis Platform for Formalin-Fixed, Paraffin-Embedded Tissues," *J. Mol. Diagnostics*, vol. 14, no. 1, pp. 22–29, Jan. 2012.
[44] M. Chen *et al.*, "Development and validation of a novel clinical fluorescence in situ hybridization assay to detect JAK2 and PD-L1 amplification: a fluorescence in situ hybridization assay for JAK2 and PD-L1 amplification," *Mod. Pathol. 2017 3011*, vol. 30, no. 11, pp. 1516–1526, Jul. 2017.
[45] E. Borazanci, S. Z. Millis, J. Kimbrough, N. Doll, D. von Hoff, and R. K. Ramanathan, "Potential actionable targets in appendiceal cancer detected by immunohistochemistry, fluorescent in situ hybridization, and mutational analysis," *J. Gastrointest. Oncol.*, vol. 8, no. 1, p. 164, 2017.
[46] F. Zakrzewski *et al.*, "Automated detection of the HER2 gene amplification status in Fluorescence in situ hybridization images for the diagnostics of cancer tissues," *Sci. Reports 2019 91*, vol. 9, no. 1, pp. 1–12, Jun. 2019.
[47] L. Annaratone *et al.*, "Quantification of HER2 and estrogen receptor heterogeneity in breast cancer by single-molecule RNA fluorescence in situ hybridization," *Oncotarget*, vol. 8, no. 12, p. 18680, Mar. 2017.
[48] T. J. Rowland, G. Dumbović, E. P. Hass, J. L. Rinn, and T. R. Cech, "Single-cell imaging reveals unexpected heterogeneity of telomerase reverse transcriptase expression across human cancer cell lines," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 116, no. 37, pp. 18488–18497, Sep. 2019.
[49] E. Lubeck, A. F. Coskun, T. Zhiyentayev, M. Ahmad, and L. Cai, "Single cell in situ RNA profiling by sequentialhybridization," *Nat. Methods*, vol. 11, no. 4, p. 360, 2014.
[50] C.-H. L. Eng *et al.*, "Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH+," *Nat. 2019 5687751*, vol. 568, no. 7751, pp. 235–239, 2019.
[51] K. H. Chen, A. N. Boettiger, J. R. Moffitt, S. Wang, and X. Zhuang, "Spatially resolved, highly multiplexed RNA profiling in single cells," *Science (80-.).*, vol. 348, no. 6233, pp. aaa6090–aaa6090, Apr. 2015.
[52] C. Xia, H. P. Babcock, J. R. Moffitt, and X. Zhuang, "Multiplexed detection of RNA using MERFISH and branched DNA amplification," *Sci. Rep.*, vol. 9, no. 1, pp. 1–13, Dec. 2019.

[53] R. Ke *et al.*, “In situ sequencing for RNA analysis in preserved tissue and cells,” *Nat. Methods*, vol. 10, no. 9, pp. 857–860, 2013.
[54] D. Gyllborg *et al.*, “Hybridization-based in situ sequencing (HybISS) for spatially resolved transcriptomics in human and mouse brain tissue,” *Nucleic Acids Res.*, vol. 48, no. 19, pp. e112–e112, Nov. 2020.
[55] T. Krzywkowski, M. Kühnemund, and M. Nilsson, “Chimeric padlock and iLock probes for increased efficiency of targeted RNA detection,” *Rna*, vol. 25, no. 1, pp. 82–89, 2019.
[56] J. Svedlund *et al.*, “Generation of in situ sequencing based OncoMaps to spatially resolve gene expression profiles of diagnostic and prognostic markers in breast cancer,” *EBioMedicine*, vol. 48, pp. 212–223, Oct. 2019.
[57] J. H. Lee, “Quantitative approaches for investigating the spatial context of gene expression.,” *Wiley Interdiscip. Rev. Syst. Biol. Med.*, vol. 9, no. 2, 2017.
[58] F. Chen, P. W. Tillberg, and E. S. Boyden, “Expansion microscopy,” *Science (80-.).*, vol. 347, no. 6221, pp. 543–548, Jan. 2015.
[59] S. Alon *et al.*, “Expansion sequencing: Spatially precise in situ transcriptomics in intact biological systems,” *Science*, vol. 371, no. 6528, Jan. 2021.
[60] X. Wang *et al.*, “Three-dimensional intact-tissue sequencing of single-cell transcriptional states,” *Science*, vol. 361, no. 6400, Jul. 2018.
[61] P. L. Ståhl *et al.*, “Visualization and analysis of gene expression in tissue sections by spatial transcriptomics,” *Science (80-.).*, vol. 353, no. 6294, pp. 78–82, Jul. 2016.
[62] S. G. Rodriques *et al.*, “Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution,” *Science (80-.).*, vol. 363, no. 6434, pp. 1463–1467, Mar. 2019.
[63] S. Vickovic *et al.*, “High-definition spatial transcriptomics for in situ tissue profiling,” *Nat. Methods*, vol. 16, no. 10, pp. 987–990, Oct. 2019.
[64] A. Andersson *et al.*, “Spatial deconvolution of HER2-positive breast cancer delineates tumor-associated cell type interactions,” *Nat. Commun. 2021 121*, vol. 12, no. 1, pp. 1–14, Oct. 2021.
[65] A. Andersson *et al.*, “Single-cell and spatial transcriptomics enables probabilistic inference of cell type topography,” *Commun. Biol. 2020 31*, vol. 3, no. 1, pp. 1–8, Oct. 2020.
[66] A. S. Booeshaghi *et al.*, “Isoform cell-type specificity in the mouse primary motor cortex,” *Nat. 2021 5987879*, vol. 598, no. 7879, pp. 195–199, Oct. 2021.

[67] L. K *et al.*, “The spatial landscape of gene expression isoforms in tissue sections,” Aug. 2020.
[68] E. Boileau *et al.*, “Full-Length Spatial Transcriptomics Reveals the Unexplored Isoform Diversity of the Myocardium Post-MI,” *Front. Genet.*, vol. 13, Jul. 2022.
[69] A. Erickson *et al.*, “Spatially resolved clonal copy number alterations in benign and malignant tissue,” *Nat. 2022 6087922*, vol. 608, no. 7922, pp. 360–367, Aug. 2022.
[70] I. Parra and B. Windle, “High resolution visual mapping of stretched DNA by fluorescent hybridization,” *Nat. Genet.*, vol. 5, no. 1, pp. 17–21, 1993.
[71] J. M. Levsky and R. H. Singer, “Fluorescence in situ hybridization: past, present and future,” *J. Cell Sci.*, vol. 116, no. 14, pp. 2833–2838, Jul. 2003.
[72] P. M. Nederlof *et al.*, “Multiple fluorescence in situ hybridization,” *Cytometry*, vol. 11, no. 1, pp. 126–131, Jan. 1990.
[73] J. G. J. Bauman, J. Wiegant, P. Borst, and P. van Duijn, “A new method for fluorescence microscopical localization of specific DNA sequences by in situ hybridization of fluorochromelabelled RNA,” *Exp. Cell Res.*, vol. 128, no. 2, pp. 485–490, 1980.
[74] Y. Takei *et al.*, “Integrated spatial genomics reveals global architecture of single nuclei,” *Nat. 2021 5907845*, vol. 590, no. 7845, pp. 344–350, Jan. 2021.
[75] A. C. Payne *et al.*, “In situ genome sequencing resolves DNA sequence and structure in intact biological samples,” *Science (80-.).*, vol. 371, no. 6532, Feb. 2021.
[76] Y. Deng *et al.*, “Spatial-CUT&Tag: Spatially resolved chromatin modification profiling at the cellular level,” *Science*, vol. 375, no. 6581, pp. 681–686, Feb. 2022.
[77] Y. Deng *et al.*, “Spatial profiling of chromatin accessibility in mouse and human tissues,” *Nat. 2022*, pp. 1–9, Aug. 2022.
[78] H. S. Kaya-Okur *et al.*, “CUT&Tag for efficient epigenomic profiling of small samples and single cells,” *Nat. Commun. 2019 101*, vol. 10, no. 1, pp. 1–10, Apr. 2019.
[79] C. A. Thornton *et al.*, “Spatially mapped single-cell chromatin accessibility,” *Nat. Commun.*, vol. 12, no. 1, pp. 1–16, 2021.
[80] G. Banik *et al.*, “High-dimensional multiplexed immunohistochemical characterization of immune contexture in human cancers,” *Methods Enzymol.*, vol. 635, pp. 1–20, Jan. 2020.
[81] T. Tsujikawa *et al.*, “Quantitative Multiplex Immunohistochemistry Reveals Myeloid-Inflamed Tumor-Immune Complexity Associated with Poor Prognosis,” *Cell Rep.*, vol. 19, no. 1, pp. 203–217, Apr.

2017.
[82] Z. E. Tóth and É. Mezey, “Simultaneous visualization of multiple antigens with tyramide signal amplification using antibodies from the same species,” *J. Histochem. Cytochem.*, vol. 55, no. 6, pp. 545–554, Jun. 2007.
[83] Y. Goltsev *et al.*, “Deep Profiling of Mouse Splenic Architecture with CODEX Multiplexed Imaging,” *Cell*, vol. 174, no. 4, pp. 968–981.e15, Aug. 2018.
[84] S. Black *et al.*, “CODEX multiplexed tissue imaging with DNA-conjugated antibodies,” *Nat. Protoc.*, vol. 16, no. 8, p. 3802, Aug. 2021.
[85] J. W. Hickey *et al.*, “Spatial mapping of protein composition and tissue organization: a primer for multiplexed antibody-based imaging,” *Nat. Methods 2021 193*, vol. 19, no. 3, pp. 284–295, Nov. 2021.
[86] Y. Hasin, M. Seldin, and A. Lusis, “Multi-omics approaches to disease,” *Genome Biol. 2017 181*, vol. 18, no. 1, pp. 1–15, May 2017.
[87] Y. Liu *et al.*, “High-Spatial-Resolution Multi-Omics Sequencing via Deterministic Barcoding in Tissue,” *Cell*, vol. 183, no. 6, pp. 1665–1681.e18, Dec. 2020.
[88] M. Stoeckius *et al.*, “Simultaneous epitope and transcriptome measurement in single cells,” *Nat. Methods 2017 149*, vol. 14, no. 9, pp. 865–868, Jul. 2017.
[89] C. R. Merritt *et al.*, “Multiplex digital spatial profiling of proteins and RNA in fixed tissue,” *Nat. Biotechnol. 2020 385*, vol. 38, no. 5, pp. 586–599, May 2020.
[90] S. Vickovic *et al.*, “SM-Omics is an automated platform for high-throughput spatial multi-omics,” *Nat. Commun. 2022 131*, vol. 13, no. 1, pp. 1–13, Feb. 2022.
[91] Z. K. Tuong *et al.*, “Resolving the immune landscape of human prostate at a single-cell level in health and cancer,” *Cell Rep.*, vol. 37, no. 12, p. 110132, Dec. 2021.
[92] M. V. Hunter, R. Moncada, J. M. Weiss, I. Yanai, and R. M. White, “Spatially resolved transcriptomics reveals the architecture of the tumor-microenvironment interface,” *Nat. Commun. 2021 121*, vol. 12, no. 1, pp. 1–16, Nov. 2021.
[93] S. Nagasawa *et al.*, “Genomic profiling reveals heterogeneous populations of ductal carcinoma in situ of the breast,” *Commun. Biol. 2021 41*, vol. 4, no. 1, pp. 1–13, Apr. 2021.
[94] K. H. Gouin *et al.*, “An N-Cadherin 2 expressing epithelial cell subpopulation predicts response to surgery, chemotherapy and immunotherapy in bladder cancer,” *Nat. Commun. 2021 121*, vol. 12, no. 1, pp. 1–14, Aug. 2021.

[95] D. B. Joseph *et al.*, "5-Alpha reductase inhibitors induce a prostate luminal to club cell transition in human benign prostatic hyperplasia," *J. Pathol.*, vol. 256, no. 4, pp. 427–441, Apr. 2022.
[96] B. A. Luca *et al.*, "Atlas of clinically distinct cell states and ecosystems across human solid tumors," *Cell*, vol. 184, no. 21, pp. 5482–5496.e28, Oct. 2021.
[97] H. Sun *et al.*, "Hypoxic microenvironment induced spatial transcriptome changes in pancreatic cancer," *Cancer Biol. Med.*, vol. 18, no. 2, pp. 616–630, May 2021.
[98] Y. Ma and X. Zhou, "Spatially informed cell-type deconvolution for spatial transcriptomics," *Nat. Biotechnol. 2022 409*, vol. 40, no. 9, pp. 1349–1359, May 2022.
[99] A. L. Ji *et al.*, "Multimodal Analysis of Composition and Spatial Architecture in Human Squamous Cell Carcinoma," *Cell*, vol. 182, no. 2, pp. 497–514.e22, Jul. 2020.
[100] S. Z. Wu *et al.*, "A single-cell and spatially resolved atlas of human breast cancers," *Nat. Genet. 2021 539*, vol. 53, no. 9, pp. 1334–1347, Sep. 2021.
[101] J. Qi *et al.*, "Single-cell and spatial analysis reveal interaction of FAP+ fibroblasts and SPP1+ macrophages in colorectal cancer," *Nat. Commun. 2022 131*, vol. 13, no. 1, pp. 1–20, Apr. 2022.
[102] M. Dhainaut *et al.*, "Spatial CRISPR genomics identifies regulators of the tumor microenvironment," *Cell*, vol. 185, no. 7, pp. 1223–1239.e20, Mar. 2022.
[103] V. M. Ravi *et al.*, "T-cell dysfunction in the glioblastoma microenvironment is mediated by myeloid cells releasing interleukin-10," *Nat. Commun. 2022 131*, vol. 13, no. 1, pp. 1–16, Feb. 2022.
[104] L. A. van de Velde *et al.*, "Neuroblastoma formation requires unconventional CD4 T cells and arginase-1–dependent myeloid cells," *Cancer Res.*, vol. 81, no. 19, pp. 5047–5059, Oct. 2021.
[105] P. Nieto *et al.*, "A single-cell tumor immune atlas for precision oncology," *Genome Res.*, vol. 31, no. 10, pp. 1913–1926, Oct. 2021.
[106] K. Thrane, H. Eriksson, J. Maaskola, J. Hansson, and J. Lundeberg, "Spatially resolved transcriptomics enables dissection of genetic heterogeneity in stage III cutaneous malignant melanoma," *Cancer Res.*, vol. 78, no. 20, pp. 5970–5979, Oct. 2018.
[107] R. Wei *et al.*, "Spatial charting of single-cell transcriptomes in tissues," *Nat. Biotechnol. 2022 408*, vol. 40, no. 8, pp. 1190–1199, Mar. 2022.
[108] R. Moncada *et al.*, "Integrating microarray-based spatial transcriptomics and single-cell RNA-seq reveals tissue architecture in pancreatic ductal adenocarcinomas," *Nat. Biotechnol. 2020 383*,

vol. 38, no. 3, pp. 333–342, Jan. 2020.
[109] T. Hara *et al.*, “Interactions between cancer cells and immune cells drive transitions to mesenchymal-like states in glioblastoma,” *Cancer Cell*, vol. 39, no. 6, pp. 779–792.e11, Jun. 2021.
[110] A. Magen *et al.*, “Intratumoral mregDC and CXCL13 T helper niches enable local differentiation of CD8 T cells following PD-1 blockade,” *bioRxiv*, p. 2022.06.22.497216, Jun. 2022.
[111] R. Tamma, T. Annese, S. Ruggieri, A. Marzullo, B. Nico, and D. Ribatti, “VEGFA and VEGFR2 RNAscope determination in gastric cancer,” *J. Mol. Histol.*, vol. 49, no. 4, pp. 429–435, Aug. 2018.
[112] C. Ruiz-Moreno *et al.*, “Harmonized single-cell landscape, intercellular crosstalk and tumor architecture of glioblastoma,” *bioRxiv*, p. 2022.08.27.505439, Aug. 2022.
[113] S. Alon *et al.*, “Expansion sequencing: Spatially precise in situ transcriptomics in intact biological systems,” *Science (80-.).*, vol. 371, no. 6528, p. eaax2656, Jan. 2021.
[114] J. Zagozewski *et al.*, “Combined MEK and JAK/STAT3 pathway inhibition effectively decreases SHH medulloblastoma tumor progression,” *Commun. Biol.*, vol. 5, no. 1, Dec. 2022.
[115] A. Sharma *et al.*, “Onco-fetal Reprogramming of Endothelial Cells Drives Immunosuppressive Macrophages in Hepatocellular Carcinoma,” *Cell*, vol. 183, no. 2, pp. 377–394.e21, Oct. 2020.
[116] K. Pelka *et al.*, “Spatially organized multicellular immune hubs in human colorectal cancer,” *Cell*, vol. 184, no. 18, pp. 4734–4752.e20, Sep. 2021.
[117] A. van Krimpen *et al.*, “Immune suppression in the tumor-draining lymph node corresponds with distant disease recurrence in patients with melanoma,” *Cancer Cell*, vol. 40, no. 8, pp. 798–799, Aug. 2022.
[118] A. Wong-Rolle *et al.*, “Spatial meta-transcriptomics reveal associations of intratumor bacteria burden with lung cancer cells showing a distinct oncogenic signature,” *J. Immunother. cancer*, vol. 10, no. 7, Jul. 2022.
[119] S. Han *et al.*, “Single-cell profiling of microenvironment components by spatial localization in pancreatic ductal adenocarcinoma,” *Theranostics*, vol. 12, no. 11, p. 4980, 2022.
[120] H. Sadeghirad *et al.*, “Dissecting Tissue Compartment-Specific Protein Signatures in Primary and Metastatic Oropharyngeal Squamous Cell Carcinomas,” *Front. Immunol.*, vol. 13, May 2022.
[121] L. Brady *et al.*, “Inter- and intra-tumor heterogeneity of metastatic prostate cancer determined by digital spatial gene expression profiling,” *Nat. Commun. 2021 121*, vol. 12, no. 1, pp. 1–16, Mar. 2021.

[122] A. Kulasinghe, J. Monkman, E. T. Shah, N. Matigian, M. N. Adams, and K. O'Byrne, "Spatial Profiling Identifies Prognostic Features of Response to Adjuvant Therapy in Triple Negative Breast Cancer (TNBC)," *Front. Oncol.*, vol. 11, Jan. 2021.
[123] K. L. McNamara *et al.*, "Spatial proteomic characterization of HER2-positive breast tumors through neoadjuvant therapy predicts response," *Nat. Cancer 2021 24*, vol. 2, no. 4, pp. 400–413, Apr. 2021.
[124] L. B. Schmitd *et al.*, "Spatial and Transcriptomic Analysis of Perineural Invasion in Oral Cancer," *Clin. Cancer Res.*, vol. 28, no. 16, pp. 3557–3572, Aug. 2022.
[125] L. Zhang *et al.*, "Single-Cell Analyses Inform Mechanisms of Myeloid-Targeted Therapies in Colon Cancer," *Cell*, vol. 181, no. 2, pp. 442–459.e29, Apr. 2020.
[126] L. H. Che *et al.*, "A single-cell atlas of liver metastases of colorectal cancer reveals reprogramming of the tumor microenvironment in response to preoperative chemotherapy," *Cell Discov. 2021 71*, vol. 7, no. 1, pp. 1–21, Sep. 2021.
[127] A. R. Pombo Antunes *et al.*, "Single-cell profiling of myeloid cells in glioblastoma across species and disease stage reveals macrophage competition and specialization," *Nat. Neurosci. 2021 244*, vol. 24, no. 4, pp. 595–610, Mar. 2021.
[128] C. M. Schürch *et al.*, "Coordinated Cellular Neighborhoods Orchestrate Antitumoral Immunity at the Colorectal Cancer Invasive Front," *Cell*, vol. 182, no. 5, pp. 1341–1359.e19, Sep. 2020.
[129] P. Mondello *et al.*, "Lack of intrafollicular memory CD4 + T cells is predictive of early clinical failure in newly diagnosed follicular lymphoma," *Blood Cancer J. 2021 117*, vol. 11, no. 7, pp. 1–11, Jul. 2021.
[130] D. Phillips *et al.*, "Immune cell topography predicts response to PD-1 blockade in cutaneous T cell lymphoma," *Nat. Commun. 2021 121*, vol. 12, no. 1, pp. 1–18, Nov. 2021.
[131] L. Ma *et al.*, "Tumor Cell Biodiversity Drives Microenvironmental Reprogramming in Liver Cancer," *Cancer Cell*, vol. 36, no. 4, pp. 418–430.e6, Oct. 2019.
[132] M. Rosa, "Fine-needle aspiration biopsy: A historical overview," *Diagn. Cytopathol.*, vol. 36, no. 11, pp. 773–775, Nov. 2008.
[133] M. R. Emmert-Buck *et al.*, "Laser capture microdissection," *Science*, vol. 274, no. 5289, pp. 998–1001, Nov. 1996.
[134] B. Domazet, G. T. MacLennan, A. Lopez-Beltran, R. Montironi, and L. Cheng, "Laser Capture Microdissection in the Genomic and Proteomic Era: Targeting the Genetic Basis of Cancer," *Int. J. Clin. Exp. Pathol.*,

vol. 1, no. 6, p. 475, Mar. 2008.
[135] R. F. Bonner, “Laser Capture Microdissection (LCM) and the Future of Molecular Pathology,” *Biomed. Opt. Spectrosc. Diagnostics / Ther. Laser Appl. (1998), Pap. JMA2*, vol. 2, p. JMA2, Mar. 1998.
[136] L. A. Liotta *et al.*, “Laser Capture Proteomics: spatial tissue molecular profiling from the bench to personalized medicine,” *Expert Rev. Proteomics*, vol. 18, no. 10, pp. 845–861, 2021.
[137] A. H. Bryce *et al.*, “Shared and unique genomic structural variants of different histological components within testicular germ cell tumours identified with mate pair sequencing,” *Sci. Rep.*, vol. 9, no. 1, Dec. 2019.
[138] P. Ellis *et al.*, “Reliable detection of somatic mutations in solid tissues by laser-capture microdissection and low-input DNA sequencing,” *Nat. Protoc. 2020 162*, vol. 16, no. 2, pp. 841–871, Dec. 2020.
[139] J. K. Tay *et al.*, “The microdissected gene expression landscape of nasopharyngeal cancer reveals vulnerabilities in FGF and noncanonical NF-κB signaling,” *Sci. Adv.*, vol. 8, no. 14, Apr. 2022.
[140] S. Tyekucheva *et al.*, “Stromal and epithelial transcriptional map of initiation progression and metastatic potential of human prostate cancer,” *Nat. Commun.*, vol. 8, no. 1, Dec. 2017.
[141] L. A. Liotta *et al.*, “Laser Capture Proteomics: spatial tissue molecular profiling from the bench to personalized medicine,” *https://doi.org/10.1080/14789450.2021.1984886*, vol. 18, no. 10, pp. 845–861, 2021.
[142] Y. Zhu *et al.*, “Spatially Resolved Proteome Mapping of Laser Capture Microdissected Tissue with Automated Sample Transfer to Nanodroplets,” *Mol. Cell. Proteomics*, vol. 17, no. 9, pp. 1864–1874, Sep. 2018.
[143] M. Schillebeeckx, A. Schrade, A. K. Löbs, M. Pihlajoki, D. B. Wilson, and R. D. Mitra, “Laser capture microdissection-reduced representation bisulfite sequencing (LCM-RRBS) maps changes in DNA methylation associated with gonadectomy-induced adrenocortical neoplasia in the mouse,” *Nucleic Acids Res.*, vol. 41, no. 11, 2013.
[144] A. Meissner, A. Gnirke, G. W. Bell, B. Ramsahoye, E. S. Lander, and R. Jaenisch, “Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis,” *Nucleic Acids Res.*, vol. 33, no. 18, p. 5868, 2005.
[145] L. Cato *et al.*, “ARv7 Represses Tumor-Suppressor Genes in Castration-Resistant Prostate Cancer,” *Cancer Cell*, vol. 35, no. 3, pp. 401-413.e6, Mar. 2019.
[146] L. Gómez-Cuadrado *et al.*, “Characterisation of the Stromal

Microenvironment in Lobular Breast Cancer," *Cancers 2022, Vol. 14, Page 904*, vol. 14, no. 4, p. 904, Feb. 2022.
[147] S. R. Chowdhuri *et al.*, "EGFR and KRAS mutation analysis in cytologic samples of lung adenocarcinoma enabled by laser capture microdissection," *Mod. Pathol.*, vol. 25, no. 4, pp. 548–555, Apr. 2012.
[148] U. Malapelle *et al.*, "EGFR and KRAS mutations detection on lung cancer liquid-based cytology: a pilot study," *J. Clin. Pathol.*, vol. 65, no. 1, pp. 87–91, Jan. 2012.
[149] J. Tay, W. K. Teo, C. K. Goh, B. C. Wu, and K. S. Loh, "Abstract 774: Microdissected gene expression profiling of recurrent nasopharyngeal carcinoma," *Cancer Res.*, vol. 82, no. 12_Supplement, pp. 774–774, Jun. 2022.
[150] M. A. Rubin *et al.*, "1Oq23.3 loss of heterozygosity is higher inlymph node-positive (PT2-3,N+) versus lymph node-negative (PT2-3,N0) prostate cancer," *Hum. Pathol.*, vol. 31, no. 4, pp. 504–508, Apr. 2000.
[151] L. Cheng, D. G. Bostwick, G. Li, S. Zhang, A. O. Vortmeyer, and Z. Zhuang, "Conserved Genetic Findings in Metastatic Bladder CancerA Possible Utility of Allelic Loss of Chromosomes 9p21 and 17p13 in Diagnosis," *Arch. Pathol. Lab. Med.*, vol. 125, no. 9, pp. 1197–1199, Sep. 2001.
[152] L. Cheng *et al.*, "Allelic Loss of the Active X Chromosome During Bladder Carcinogenesis," *Arch. Pathol. Lab. Med.*, vol. 128, no. 2, pp. 187–190, Feb. 2004.
[153] P. Bertheau *et al.*, "Allelic Loss Detection in Inflammatory Breast Cancer: Improvement with Laser Microdissection," *Lab. Investig. 2001 8110*, vol. 81, no. 10, pp. 1397–1402, Oct. 2001.
[154] X. Zhang, I. Leav, M. P. Revelo, R. Deka, and M. Medvedovic, "Deletion Hotspots in AMACR Promoter CpG Island Are cis-Regulatory Elements Controlling the Gene Expression in the Colon," *PLoS Genet*, vol. 5, no. 1, p. 1000334, 2009.
[155] P. Wild, R. Knuechel, W. Dietmaier, F. Hofstaedter, and A. Hartmann, "Laser Microdissection and Microsatellite Analyses of Breast Cancer Reveal a High Degree of Tumor Heterogeneity," *Pathobiology*, vol. 68, no. 4–5, pp. 180–190, 2000.
[156] L. Cheng *et al.*, "Precise microdissection of human bladder carcinomas reveals divergent tumor subclones in the same tumor," *Cancer*, vol. 94, no. 1, pp. 104–110, Jan. 2002.
[157] T. D. Jones, J. N. Eble, M. Wang, G. T. MacLennan, S. Jain, and L. Cheng, "Clonal divergence and genetic heterogeneity in clear cell renal cell carcinomas with sarcomatoid transformation," *Cancer*, vol.

104, no. 6, pp. 1195–1203, Sep. 2005.
[158] T. M. Katona, T. D. Jones, M. Wang, J. N. Eble, S. D. Billings, and L. Cheng, “Genetically heterogeneous and clonally unrelated metastases may arise in patients with cutaneous melanoma,” *Am. J. Surg. Pathol.*, vol. 31, no. 7, pp. 1029–1037, Jul. 2007.
[159] S. Olafsson and C. A. Anderson, “Somatic mutations provide important and unique insights into the biology of complex diseases,” *Trends Genet.*, vol. 37, no. 10, pp. 872–881, Oct. 2021.
[160] S. M. Cowherd, V. A. Espina, E. F. Petricoin, and L. A. Liotta, “Proteomic Analysis of Human Breast Cancer Tissue with Laser-Capture Microdissection and Reverse-Phase Protein Microarrays,” *Clin. Breast Cancer*, vol. 5, no. 5, pp. 385–392, Dec. 2004.
[161] R. Buckanovich *et al.*, “Tumor Vascular Proteins As Biomarkers in Ovarian Cancer Article in Journal of Clinical Oncology ·,” *J Clin Oncol*, vol. 25, pp. 852–861, 2007.
[162] K. H. B. Lam *et al.*, “Topographic mapping of the glioblastoma proteome reveals a triple-axis model of intra-tumoral heterogeneity,” *Nat. Commun. 2022 131*, vol. 13, no. 1, pp. 1–14, Jan. 2022.
[163] S. A. Selamat *et al.*, “Genome-scale analysis of DNA methylation in lung adenocarcinoma and integration with mRNA expression,” *Genome Res.*, vol. 22, no. 7, p. 1197, Jul. 2012.
[164] M. Schillebeeckx, A. Schrade, A. K. Löbs, M. Pihlajoki, D. B. Wilson, and R. D. Mitra, “Laser capture microdissection–reduced representation bisulfite sequencing (LCM-RRBS) maps changes in DNA methylation associated with gonadectomy-induced adrenocortical neoplasia in the mouse,” *Nucleic Acids Res.*, vol. 41, no. 11, pp. e116–e116, Jun. 2013.
[165] L. Zhao, X. Wu, J. Zheng, and D. Dong, “DNA methylome profiling of circulating tumor cells in lung cancer at single base-pair resolution,” *Oncogene 2021 4010*, vol. 40, no. 10, pp. 1884–1895, Feb. 2021.
[166] B. Jovanovic *et al.*, “A randomized phase II neoadjuvant study of cisplatin, paclitaxel with or without everolimus in patients with stage II/III triple-negative breast cancer (TNBC): Responses and long-term outcome correlated with increased frequency of DNA damage response gene mutations, TNBC subtype, AR status, and Ki67,” *Clin. Cancer Res.*, vol. 23, no. 15, pp. 4035–4045, Aug. 2017.
[167] T. Chen, C. Cao, J. Zhang, A. Streets, T. Li, and Y. Huang, “Histologically resolved multiomics enables precise molecular profiling of human intratumor heterogeneity,” *PLOS Biol.*, vol. 20, no. 7, p. e3001699, Jul. 2022.
[168] K. Krysan *et al.*, “The immune contexture associates with the

genomic landscape in lung adenomatous premalignancy," *Cancer Res.*, vol. 79, no. 19, pp. 5022–5033, Oct. 2019.
[169] Z. Zhu *et al.*, "Genome profiles of pathologist-defined cell clusters by multiregional LCM and G&T-seq in one triple-negative breast cancer patient," *Cell reports. Med.*, vol. 2, no. 10, Oct. 2021.
[170] N. Nitta *et al.*, "Raman image-activated cell sorting," *Nat. Commun. 2020 111*, vol. 11, no. 1, pp. 1–16, Jul. 2020.
[171] N. Nitta *et al.*, "Intelligent Image-Activated Cell Sorting," *Cell*, vol. 175, no. 1, pp. 266–276.e13, Sep. 2018.
[172] S. Kim *et al.*, "PHLI-seq: constructing and visualizing cancer genomic maps in 3D by phenotype-based high-throughput laser-aided isolation and sequencing," *Genome Biol.*, vol. 19, no. 1, p. 158, Oct. 2018.
[173] A. C. Lee *et al.*, "Spatial epitranscriptomics reveals A-to-I editome specific to cancer stem cell microniches," *Nat. Commun. 2022 131*, vol. 13, no. 1, pp. 1–12, May 2022.
[174] O. Kim *et al.*, "Whole Genome Sequencing of Single Circulating Tumor Cells Isolated by Applying a Pulsed Laser to Cell-Capturing Microstructures," *Small*, vol. 15, no. 37, p. 1902607, Sep. 2019.
[175] M. L. Metzker, "Sequencing technologies - the next generation.," *Nat. Rev. Genet.*, vol. 11, no. 1, pp. 31–46, 2010.
[176] V. Marx, "Method of the Year: spatially resolved transcriptomics," *Nat. Methods 2021 181*, vol. 18, no. 1, pp. 9–14, Jan. 2021.
[177] H. M. Davey and D. B. Kell, "Flow cytometry and cell sorting of heterogeneous microbial populations: the importance of single-cell analyses.," *Microbiol. Rev.*, vol. 60, no. 4, p. 641, Dec. 1996.
[178] M. R. Emmert-Buck *et al.*, "Laser capture microdissection," *Science (80-.).*, vol. 274, no. 5289, pp. 998–1001, Nov. 1996.
[179] S. Kim *et al.*, "PHLI-seq: constructing and visualizing cancer genomic maps in 3D by phenotype-based high-throughput laser-aided isolation and sequencing," *Genome Biol.*, vol. 19, no. 1, pp. 1–17, Oct. 2018.
[180] H. E. Muñoz, J. Che, J. E. Kong, and D. Di Carlo, "Advances in the production and handling of encoded microparticles," *Lab Chip*, vol. 14, no. 13, pp. 2212–2216, Jun. 2014.
[181] S. W. Song, "One-Step Generation of a Drug-Releasing Microarray for High-Throughput Small-Volume Bioassays", *Springer Theses,* 2019.
[182] N. M. Karabacak *et al.*, "Microfluidic, marker-free isolation of circulating tumor cells from blood samples," *Nat. Protoc. 2013 93*, vol. 9, no. 3, pp. 694–710, Feb. 2014.
[183] W. J. Allard *et al.*, "Tumor Cells Circulate in the Peripheral Blood of

All Major Carcinomas but not in Healthy Subjects or Patients With Nonmalignant Diseases," *Clin. Cancer Res.*, vol. 10, no. 20, pp. 6897–6904, Oct. 2004.
[184] C. McQuin *et al.*, "CellProfiler 3.0: Next-generation image processing for biology," *PLOS Biol.*, vol. 16, no. 7, p. e2005970, Jul. 2018.
[185] S. Picelli, O. R. Faridani, Å. K. Björklund, G. Winberg, S. Sagasser, and R. Sandberg, "Full-length RNA-seq from single cells using Smart-seq2," *Nat. Protoc. 2013 91*, vol. 9, no. 1, pp. 171–181, Jan. 2014.
[186] S. M. Lewis *et al.*, "Spatial omics and multiplexed imaging to explore cancer biology," *Nat. Methods*, vol. 18, no. 9, pp. 997–1012, 2021.
[187] S. Bae, K. J. Na, J. Koh, D. S. Lee, H. Choi, and Y. T. Kim, "CellDART: cell type inference by domain adaptation of single-cell and spatial transcriptomic data," *Nucleic Acids Res.*, vol. 50, no. 10, pp. e57–e57, Jun. 2022.

국문 초록

종양은 종양세포만이 원인이 되는 질병이 아닌, 면역세포와 혈관 구조, 표피 세포 등 다양한 세포들이 공동체를 이뤄 질병을 발전시키는 복합적인 질병이다. 이를 분석하기 위하여 차세대 시퀀싱 (NGS), 단일 세포 분석, 유세포 분석 등 다양한 기술이 활용되어 왔다. 특히 RNA sequencing 을 포함한 단일 세포 분석은 종양 내에서 성장, 전이, 진화, 약물의 내성과 관련된 타겟을 발굴하고 이를 조절할 수 있는 인자들을 개발하는 등 종양 연구의 혁신을 이끌었다. 그러나 이러한 접근 방법은 종양 미세 환경 내에서 세포가 존재하고 있는 위치 정보가 손실 되기 때문에 종양의 온전한 그림을 제공하지 못한다. 최근에는 종양 내 이형적으로 분포하고 있는 유전 물질들을 이해하면 종양의 이형성과 종양 미세환경의 탐구와 암의 진화 및 발전 메커니즘을 이해하여 이전에는 발견하지 못한 새로운 제약, 진단 타겟을 발굴 할 수 있을 것이라는 관점들이 제시되고 있다.

다중의 형광 프로브를 활용한 방식으로 DNA, RNA, 단백질 등의 위치를 표적하여 종양의 전체적인 그림을 파악하는 기술들이 최근

폭발적으로 개발되고 있고, 현장 바코딩 방식으로 위치 별로 발현하고 있는 유전 물질을 표적하여 정량화 할 수 있는 방법도 상용화 되어 종양에서 새로운 발견을 이끌어내고 있다. 그러나 위의 제시된 기술들 모두 *De novo*로 새로운 제약 진단 타겟을 발굴하는 데에는 한계가 존재한다.

본 논문에서는 공간 오믹스를 가능하게 하는 세포 분리 분석기를 개발하여 관심영역의 세포를 분리하고 이후 DNA, RNA, 단백질 등을 분석하는 어세이들과 연결하여 위치정보를 포함하면서도 높은 정보량으로 세포를 분석하고, 이전까지 발굴하지 못하였던 새로운 진단 제약 타겟을 탐색할 수 있는 기술에 대하여 설명한다. 레이저를 조사하면 희생층이 승화되며 조사 영역에 위치한 세포들을 회수할 수 있고, 현존하는 생화학 어세이를 수행 할 수 있다. 자동으로 이미지 프로세싱을 통하여 원하는 세포들을 특정할 수 있는 프로그램을 개발하였고, 이를 세포주와 임상 샘플에 적용하고 이후 생화학 어세이 적용을 입증하였다. 항원 항체 반응을 이용한 염색 방법 이외에도 높은 정보량으로 유전체 지도를 그리는 염색법 등 다양한 염색 이미지에 적용했으며 본 기술을 활용하여 향후 응용할 수 있는 연구를 제안하였다.

주요어 : 세포분리분석기, 멀티오믹스, 공간유전학, 이미지 프로세싱
학번 : 2018-24629