

# 단일세포 유전자 발현 및 크로마틴 접근성 데이터의 통합

- 10x Genomics사의 Chromium Single Cell Multiome ATAC + Gene Expression을 활용하여 동일한 단일 세포로부터 다양한 오믹스 데이터를 측정 가능
- NovaSeq™ X, NovaSeq 6000, NextSeq™ 2000, NextSeq 1000 또는 NextSeq 550 시스템에서 페어링된 ATAC-Seq 라이브러리와 3' Gene Expression 라이브러리의 시퀀싱
- 유전자 조절에 대한 총체적 접근을 위해 정확한 단일세포 전사체와 후성유전체 분석 결과의 연계 지원



## 소개

단일세포(Single-cell) 분석은 복합적인 세포군 내 이질성(heterogeneity)을 확인하여 세포 유형을 구분하고 동적인 세포 상태를 파악하는 데 도움을 줄 수 있습니다. 차세대 시퀀싱(Next-generation sequencing, NGS) 기술은 혁신적인 assay를 통해 수백에서 수만 개의 세포로부터 유전체(genome), 전사체(transcriptome), 후성유전체(epigenome), 단백체(proteome)를 단일세포 해상도로 연구할 수 있도록 해 줍니다. 연구자는 이러한 상보적인 지표들(phenotype)에 대한 이해의 폭을 넓힐 수 있습니다.

대다수의 멀티오믹스 데이터 분석 방법은 정교한 알고리즘을 사용하고 다양한 데이터 표현 형식(modality) 간의 상관 관계를 추론함으로써 여러 실험으로부터 얻은 데이터를 연계합니다.<sup>1</sup> 반면에 단일세포 시퀀싱 워크플로우는 동일한 세포에서 세포의 몇 가지 특징을 파악한 후 올리고뉴클레오타이드(oligonucleotide) 바코드를 사용하여 분석 결과를 연계하는 방식을 적용해 멀티오믹스 데이터 세트를 assay 내에 구축합니다.

본 Technical Note는 단일 세포로부터 전사체(3' 유전자 발현(gene expression) 분석 이용) 및 후성유전체(ATAC-Seq (assay for transposase-accessible chromatin with sequencing) 이용)를 동시에 프로파일링하는 프로토콜을 소개합니다. 연구자는 동일한 세포에서 단일세포의 유전자 발현과 크로마틴 접근성(chromatin accessibility) 데이터를 정밀하게 연계함으로써 어떻게 유전자가 다양한 세포 유형에 걸쳐 발현되고 조절되는지를 밝혀낼 수 있습니다.

## 프로토콜의 개요

아래의 단일세포 멀티오믹스 실험은 라이브러리 준비, 시퀀싱, 분석의 단계로 구성된 직관적인 워크플로우를 통해 진행됩니다 (그림 1). 이 프로토콜에는 10x Genomics의 Chromium Next GEM Single Cell Multiome ATAC + Gene Expression과 Illumina의 입증된 시퀀싱 기술이 적용되었습니다. 핵 현탁액(Nuclei suspension)을 분리한 후, Chromium X/iX와 시약을 이용하여 시퀀싱에 바로 사용할 수 있는 바코드화된 2개의 멀티오믹(multiome) 라이브러리(Single Cell ATAC-Seq 라이브러리 1개와 Single Cell Gene Expression 라이브러리 1개)를 생성합니다.<sup>2</sup> 이렇게 페어링된 멀티오믹 라이브러리를 NovaSeq X, NovaSeq 6000, NextSeq 2000, NextSeq 1000 또는 NextSeq 550 시스템과 같은 대용량 Illumina 시퀀싱 시스템으로 시퀀싱합니다. Cell Ranger ARC 파이프라인 또는 Illumina DRAGEN™ Single-Cell 앱을 이용한 데이터 분석은 개별 세포에 대한 ATAC-Seq 결과와 유전자 발현 분석 결과를 해당 세포의 바코드를 이용해 연계합니다. Loupe Browser 소프트웨어는 연구자가 손쉽게 단일세포 멀티오믹스 데이터를 시각화하고 탐색할 수 있도록 해 줍니다.

## 샘플 준비

샘플 준비 단계는 세포 배양체(culture), 1차 세포(primary cell) 또는 신선한 조직이나 냉동된 조직에서 핵 현탁액을 분리하는 작업으로 시작됩니다. 본 assay는 다양한 세포주(cell line), 말초 혈액 단핵 세포(peripheral blood mononuclear cell, PBMC), 신선한 또는 냉동된 마우스 배아의 뇌, 냉동된 인간의 뇌, 냉동된 림프절 종양 샘플이 사용된 실험을 통해 검증되었으며, 다른 유형의 샘플에도 적용이 가능합니다.<sup>3</sup>

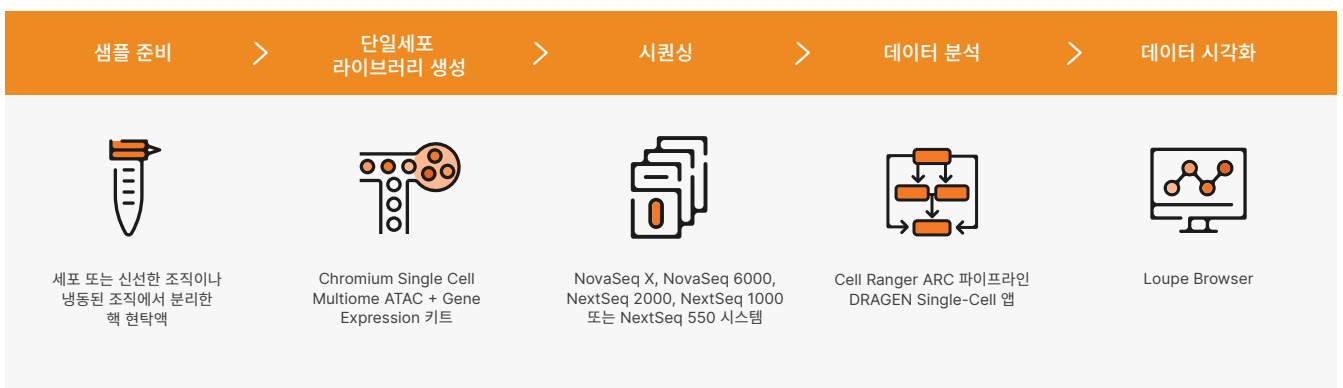


그림 1: 단일세포 멀티오믹스 실험 워크플로우 — 핵 현탁액 분리 후 Chromium X/iX로 미세유체 기반의 단일세포 분할과 바코드화를 진행함. 이렇게 생성된 2개의 멀티오믹 라이브러리(Single Cell ATAC-Seq 라이브러리 1개와 Single Cell Gene Expression 라이브러리 1개)를 Illumina 시스템으로 시퀀싱함. 데이터는 Cell Ranger ARC 파이프라인 또는 Illumina DRAGEN Single-Cell 앱(BaseSpace™ Sequence Hub 또는 NextSeq 1000 및 NextSeq 2000 시스템 내 제공)을 사용해 분석한 후 Loupe Browser 소프트웨어로 시각화함.

최적의 결과를 얻기 위해서는 양질의 핵을 사용하는 것이 매우 중요합니다. 10x Genomics 웹사이트의 Support 페이지는 다양한 유형의 샘플을 통해 입증된 핵 분리 프로토콜을 제공하고 있습니다.<sup>4</sup> 이러한 입증된 프로토콜은 세포 현탁액 및 조직의 냉동 방법, DNase 처리와 같은 클린업 방법, 형광 활성화 세포 분류(fluorescence-activated cell sorting, FACS)를 통한 과립구(granulocyte) 세포 제거 방법, FACS를 통한 핵 분류 방법을 안내합니다.<sup>3</sup>

### 단일세포 라이브러리 생성

핵이 분리가 되면, Chromium Next GEM Single Cell Multiome ATAC + Gene Expression 키트(그림 2)<sup>4</sup>를 이용하여 라이브러리를 만들 수 있습니다. 과활성 트랜스포사제 엔자임(Hyperactive transposase enzyme)으로 대량의 핵을 처리하여 노출된 DNA<sup>5</sup>를 절단하고 시퀀싱 어댑터(adapter)를 삽입합니다. 전위된 핵(Transposed nuclei)을 미세유체 칩(microfluidic chip)에 넣고 Chromium X/iX를 실행합니다. Chromium X/iX는 각각의 핵을 고유의 바코드가 부여된 1개의 겔 비드를 가진 액적(droplet)으로 분할합니다. 이러한 액적을 GEM(gel bead-in-emulsion)이라 합니다. 이제 동일한 핵에서 얻은 mRNA와 전위된 DNA 절편에 바코드를 부착하기 위해 GEM을 배양합니다. 이 단계에서 ATAC-Seq 결과와 유전자 발현 분석 결과가 연계됩니다.

배양이 끝나면 GEM을 부수어서 풀링된 조각을 회수하고 정제합니다. 정제된 조각은 갭(gap)을 채우고 바코드화된 전위 DNA와 cDNA 절편을 최대한 회수하기 위해 예비 증폭(pre-amplification) 중합효소 연쇄 반응(polymerase chain reaction, PCR) 과정을 거치게 됩니다. 그다음 예비 증폭

산물은 Gene Expression 라이브러리 구축 시 cDNA 증폭, 그리고 ATAC-Seq 라이브러리 구축에 사용됩니다.<sup>3</sup> 마지막으로 바코드화된 Single Cell Multiome Gene Expression 라이브러리 및 Single Cell Multiome ATAC-Seq 라이브러리를 Illumina의 NGS 시스템으로 시퀀싱하면 됩니다.

### Illumina 시스템을 사용한 시퀀싱

페어링된 멀티오믹 라이브러리는 이 애플리케이션에 요구되는 시퀀싱 아웃풋을 충족하는 NovaSeq X, NovaSeq 6000, NextSeq 2000, NextSeq 1000 또는 NextSeq 550 시스템으로 시퀀싱하는 것을 권장합니다(표 1). 리드(Read) 구성은 사용된 라이브러리의 유형에 따라 결정됩니다(표 2, 표 3).

멀티오믹 ATAC-Seq 라이브러리에는 1% PhiX spike-in이 권장됩니다. Spike-in은 시퀀싱의 다양성(diversity)을 적절히 보장하여 양질의 시퀀싱 결과를 얻을 수 있도록 해 줍니다.<sup>7</sup> Single Cell Multiome Gene Expression 라이브러리<sup>8</sup>와 Single Cell Multiome ATAC-Seq 라이브러리<sup>9</sup>에 대한 예상되는 시퀀싱 매트릭스는 10x Genomics 웹사이트의 Support 페이지에서 확인하실 수 있습니다.

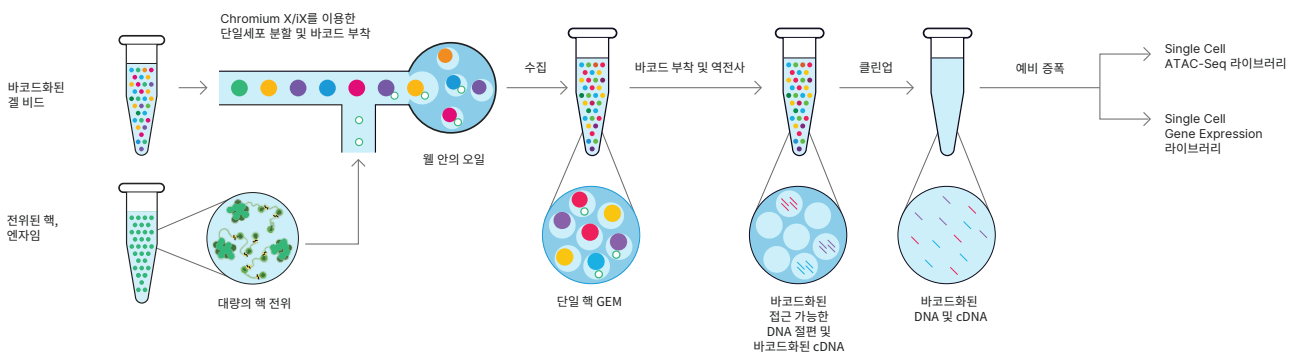


그림 2: 동일한 세포로 3' Gene Expression 라이브러리 및 ATAC-Seq 라이브러리 생성 — 단일 핵 현탁액(Single-nuclei suspension) 분리 후 대량의 핵을 전위하여 DNA 절편과 mRNA의 3' 말단이 바코드화되는 GEM 안에 각각의 핵을 포획함. 이후 GEM은 분리 및 풀링, 클린업, 예비 증폭 그리고 라이브러리 구축 단계를 거치며, 이 과정을 통해 유전자 발현 프로파일과 열린 크로마틴 프로ファイルを 동일한 세포와 확실히 연계하기 위해 각 샘플로부터 2개의 상보적인 라이브러리를 생성함.<sup>2</sup>

표 1: Illumina 시퀀싱 시스템으로 Chromium Single Cell Multiome Assay 수행 시 샘플 처리량 예시

라이브러리 유형	핵당 최소 리드 페어 수 <sup>a</sup>	샘플당 핵 수	런당 샘플 수								
			NextSeq 2000 <sup>d</sup>		NovaSeq 6000				NovaSeq X		
			P2 <sup>e</sup>	P3	SP	S1	S2	S4	1.5B	10B	25B
Multioime 3' Gene Expression	20,000 <sup>b</sup>	5,000	4	11	8	16	40	100 <sup>f</sup>	15	100	250
Multioime ATAC-Seq	25,000 <sup>b,c</sup>	5,000	4	8	6	12	32	80	12	80	200

- a. 최소 권장 리드 수(10x Genomics 제공)<sup>7</sup>
- b. 요구되는 성능이나 애플리케이션에 따라 시퀀싱 뎁스(depth) 조정이 필요함. Cell Ranger ARC 런(run) 요약 내용 중 시퀀싱 포화(saturation) 지표와 곡선을 참고하여 특정 유형의 샘플에 맞게 시퀀싱 뎁스 최적화가 가능함.
- c. 개별 리드 수: 50,000(Read 1: 25,000, Read 2: 25,000)
- d. NextSeq 550 시스템 사용 시 Multioime 3' Gene Expression 라이브러리는 런당 샘플 4개, Multioime ATAC-Seq 라이브러리는 런당 샘플 3개
- e. NextSeq 1000 시스템에서도 동일한 샘플 처리량을 지원하는 P2 플로우 셀 이용이 가능함.
- f. 샘플 100개 로딩 시 개별 레인(lane) 로딩이 필요함.

표 2: Chromium Single Cell Multiome 라이브러리의 권장 리드 구성

	Multioime 3' Gene Expression 라이브러리				Multioime ATAC-Seq 라이브러리			
	Read 1	i7 Index	i5 Index	Read 2	Read 1	i7 Index	i5 Index	Read 2
용도	세포 바코드 및 UMI	샘플 인덱스	샘플 인덱스	cDNA insert	전위 DNA	샘플 인덱스	세포 바코드	전위 DNA
리드 길이 <sup>a</sup>	28 bp	10 bp	10 bp	90 bp	50 bp <sup>b</sup>	8 bp	24 bp <sup>c</sup>	49 bp <sup>b</sup>

- a. 전사물(Transcript) 리드 길이가 짧을수록 전사체 정렬(alignment) 비율도 감소할 가능성이 있음. 세포 바코드, 고유한 분자 식별자(unique molecular identifier, UMI), 샘플 인덱스 리드는 명시된 길이보다 짧아서는 안 됨. 단, 권장된 길이보다 긴 리드는 사용이 가능함. Cell Ranger ARC를 통해 세포 바코드 또는 UMI 리드에 존재하는 추가적인 염기(base)가 자동으로 배제됨.
- b. 시퀀싱 길이는 시퀀싱 키트에 따라 조정 가능하나 30 bp 미만이어서는 안 됨.
- c. NextSeq 550 시스템과 같이 24 bp i5 Index Read를 지원하지 않는 시퀀싱 시스템에는 커스텀 레시피(custom recipe)가 필요함.<sup>10</sup>

표 3: Illumina 시퀀싱 시스템으로 Chromium Single Cell Multiome Assay 수행 시 분석 옵션

시스템	커스텀 레시피 필요 여부	FASTQ 파일 생성 옵션	멀티오믹 분석 옵션
NextSeq 500 및 NextSeq 550 시스템	필요(Multioime ATAC-Seq 라이브러리) 불필요(Multioime 3' Gene Expression 라이브러리)	기기 내 bcl2fastq bcl-convert Cell Ranger ARC mkfastq	Cell Ranger ARC DRAGEN Single-Cell 앱
NextSeq 1000 및 NextSeq 2000 시스템	불필요	기기 내 bcl2fastq bcl-convert Cell Ranger ARC mkfastq	Cell Ranger ARC DRAGEN Single-Cell 앱
NovaSeq 6000 시스템	불필요	bcl2fastq bcl-convert Cell Ranger ARC mkfastq	Cell Ranger ARC DRAGEN Single-Cell 앱
NovaSeq X 시리즈 <sup>a</sup>	불필요	bcl2fastq bcl-convert	Cell Ranger ARC DRAGEN Single-Cell 앱

a. NovaSeq X 시리즈에서 생성된 시퀀싱 데이터는 Illumina의 디멀티플렉싱(demultiplexing) 파이프라인(예: bcl2fastq 또는 bcl-convert)을 통해 FASTQ 파일로 변환됨.

## 데이터 분석 및 시각화

시퀀싱이 끝나면 10x Genomics의 Cell Ranger ARC 분석 파이프라인이 단일세포 내 열린 크로마틴 영역을 식별함과 동시에 전사물의 수와 피크 접근성(peak accessibility)을 측정합니다. 동일한 세포에서 ATAC-Seq 및 전사체 측정이 이루어지므로 크로마틴 접근성과 유전자 발현의 판독 결과를 바로 연계시킬 수 있습니다.

Cell Ranger ARC 소프트웨어는 프로필이 유사한 세포 클러스터(cluster)도 식별합니다. 분석 파이프라인은 10x Genomics의 Loupe Browser 시각화 소프트웨어나 R 또는 Python과 같은 서드 파티 도구를 활용한 추가적인 데이터 분석에도 사용이 가능한 파일과 QC(quality control, 품질 관리) 정보<sup>1)</sup>를 생성합니다.

Illumina의 DRAGEN Single-Cell 앱은 클라우드, 온프레미스 서버(on-premise server) 또는 NextSeq 1000 및 NextSeq 2000 시스템 내(onboard)에서 사용 가능하며 단일세포 멀티오믹스 데이터 분석 도구도 제공합니다. 이 강력한 분석 파이프라인은 QC 지표와 Scanpy, AnnData, Seurat 등 널리 사용되는 단일세포 분석 도구와 호환되는 세포별 유전자 발현 행렬(cell-by-gene expression matrix)을 생성합니다.

## 주요 데이터

전사체학적 분석과 후성유전체학적 분석을 연계하면 세포 유형과 상태의 특징을 심층적으로 이해할 수 있습니다. 예를 들어, 동일한 세포 유형에서 전사 인자(transcription factor) 발현을 모티프(motif) 접근성과 비교함으로써 차등 유전자 발현(differential gene expression)의 드라이버(driver)를 식별하고 더 정교한 유전자 조절 네트워크(gene regulatory network)를 구축할 수 있습니다(그림 3).<sup>2)</sup> 뿐만 아니라 Loupe Browser를 활용하면 인접한 유전자 발현과 관련된 열린 크로마틴 피크 간의 연관성도 시각화할 수 있습니다(그림 4).<sup>2)</sup>

## 전문가 지원

Illumina와 10x Genomics의 기술지원팀은 Chromium Single Cell Multiome ATAC + Gene Expression 워크플로우 전반에 걸친 지원을 하고 있습니다. Assay 및 분석 관련 질문은 10x Genomics의 기술지원팀(support@10xgenomics.com)으로, 시퀀싱 관련 질문은 Illumina의 기술지원팀(techsupport@illumina.com)으로 문의 가능합니다. 또한 Illumina와 10x Genomics의 기술지원팀은 복잡한 문제를 해결하는 데에도 도움을 드릴 수 있습니다.

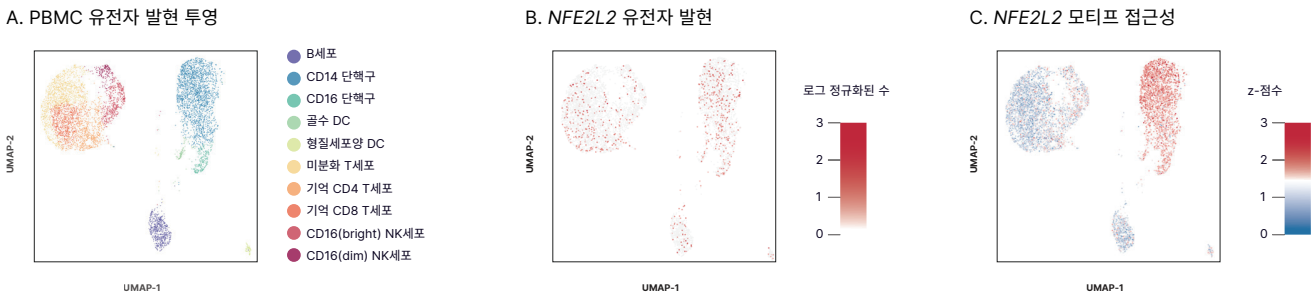


그림 3: 유전자 발현 및 크로마틴 상태에 대한 상보적인 단일세포 데이터 — 건강한 PBMC에서 추출한 핵을 Chromium Single Cell Multiome ATAC + Gene Expression으로 처리 후 NovaSeq 6000 시스템으로 라이브러리를 시퀀싱함. (A) 유전자 발현 데이터를 이용해 7,273개의 핵에 대한 클러스터 분석을 수행한 결과로, 세포군의 주석(annotation)은 기존의 표지 유전자(marker gene)에 기반함. (B) 전사 인자 NFE2L2의 발현은 모든 세포 유형에 걸쳐 관찰됨. (C) 동일한 세포로부터 얻은 ATAC-Seq 데이터로 측정된 NFE2L2 모티프 접근성은 단핵구(monocyte)군에 국한됨.<sup>2)</sup>

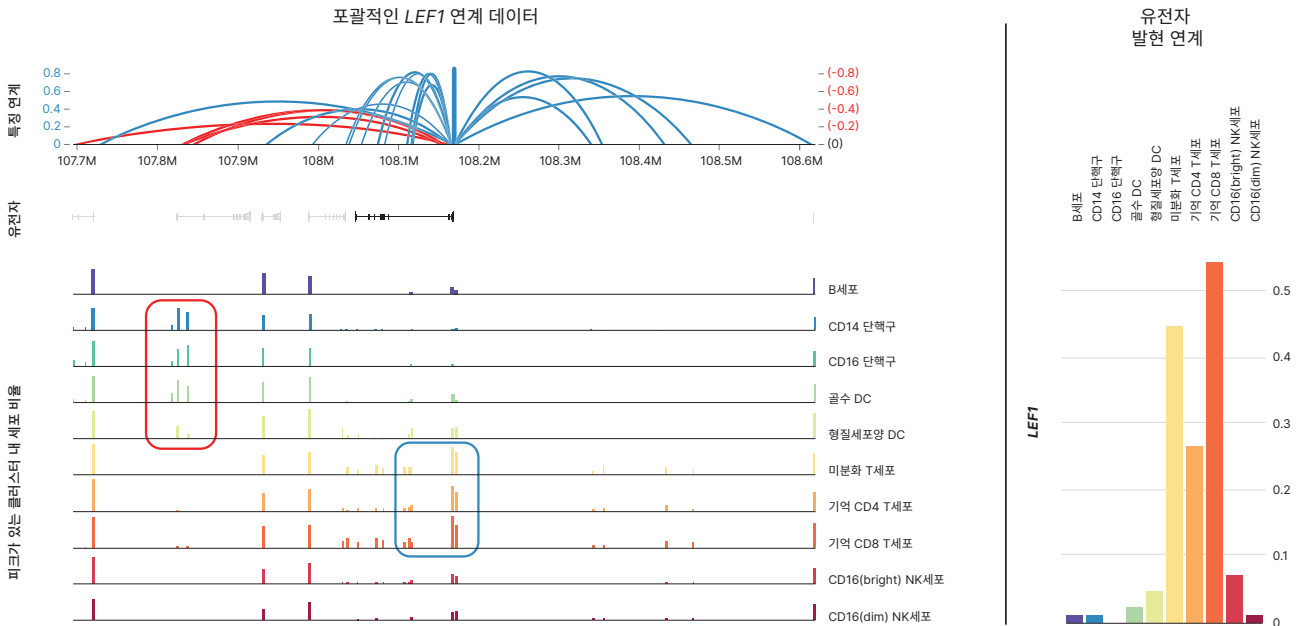


그림 4: 관심 유전자와 직접 연계된 추정 조절 요소의 식별 — 포괄적인 *LEF1* 연계 데이터는 앞서 그림 3에서 분석한 PBMC 핵 7,273개의 1 Mb 영역 전반에 대해 *LEF1* 유전자 발현과 양의 상관관계가 있는 열린 크로마틴 피크(파란색 곡선)와 음의 상관관계가 있는 열린 크로마틴 피크(빨간색 곡선)를 나타냄. *LEF1* 발현 수준과 열린 크로마틴 피크는 세포 유형에 따라 색으로 구분됨. *LEF1*의 세포 유형 특이적인 발현은 특히 미분화 T세포 및 기억 T세포(naïve & memory T-cell; 파란색 상자)에서 발현이 증가(enriched)된 *LEF1* 프로모터(promoter) 근처에 있는 연계된 열린 크로마틴 영역과 상관관계를 보임. 단핵구나 골수계 수지상 세포(myeloid dendritic cell)와 같이 *LEF1* 발현량이 적은 세포의 열린 크로마틴 영역은 각각 최소 수백 Kb의 거리에 존재하며 발현이 억제되어 있을 가능성이 있음(빨간색 상자).<sup>2</sup>

## 요약

본 단일세포 멀티오믹스 프로토콜은 단일 세포의 유전자 발현과 크로마틴 접근성을 동시에 프로파일링하는 데 적합한 프로토콜입니다. 통합된 데이터 세트는 질병 유무에 따른 유전자 발현의 차이를 비롯한 유전자 조절을 유도하는 세포 기전을 밝혀내는 데 유용합니다.

## 상세 정보

단일세포 시퀀싱

Chromium Single Cell Multiome ATAC + Gene Expression

NovaSeq X 시리즈를 활용한 정확도 높은 NGS

## 참고 문헌

1. Stuart T, Butler A, Hoffman P, et al. [Comprehensive Integration of Single-Cell Data](#). *Cell*. 2019;177(7):1888-1902.e21. doi:10.1016/j.cell.2019.05.031
2. 10x Genomics. [Simultaneous profiling of the transcriptome and epigenome from the same cell](#). Accessed October 25, 2023.
3. 10x Genomics. Single Cell Multiome ATAC + Gene Expression Support. [support.10xgenomics.com/single-cell-multiome-atac-gex](https://support.10xgenomics.com/single-cell-multiome-atac-gex). Accessed October 25, 2023.
4. 10x Genomics. Single Cell Multiome ATAC + Gene Expression Sample Prep Demonstrated Protocols. [support.10xgenomics.com/single-cell-multiome-atac-gex/sample-prep](https://support.10xgenomics.com/single-cell-multiome-atac-gex/sample-prep). Accessed October 25, 2023.
5. Buenrostro J, Wu B, Chang H, Greenleaf W. [ATAC-seq: a method for assaying chromatin accessibility genome-wide](#). *Curr Protoc Mol Biol*. 2015;109:21.29.1-21.29-9. doi:10.1002/0471142727.mb2129s109
6. 10x Genomics. [Specifying Input FASTQ Files for cellranger-arc count](#). Accessed October 10, 2023.
7. 10x Genomics. [Sequencing Requirements for Single Cell Multiome ATAC + Gene Expression](#). Accessed October 10, 2023.
8. 10x Genomics. [Sequencing Metrics & Base Composition of Single Cell 3' v3.1 Dual Index Libraries](#). Accessed October 25, 2023.
9. 10x Genomics. [Sequencing Metrics & Base Composition of Single Cell Multiome ATAC Libraries](#). Accessed October 10, 2023.
10. 10x Genomics. [Why do I need a custom recipe when sequencing Multiome ATAC libraries on the NextSeq?](#) Accessed October 25, 2021.
11. 10x Genomics. [Interpreting Cell Ranger ARC Web Summary Files for Single Cell Multiome ATAC + Gene Expression Assay](#). Accessed October 25, 2023.



무료 전화(한국) 080-234-5300  
techsupport@illumina.com | www.illumina.com

© 2024 Illumina, Inc. All rights reserved. 모든 상표는 Illumina, Inc. 또는 각 소유주의 자산입니다.  
특정 상표 정보는 [www.illumina.com/company/legal.html](http://www.illumina.com/company/legal.html)을 참조하십시오.  
M-KR-00092 v1.1 KOR