

# Illumina DRAGEN™ Secondary Analysis를 랩 인프라에 통합하는 방법

엔드투엔드 연구  
워크플로우를 위한  
DRAGEN Secondary  
Analysis 배포 옵션



## 소개

차세대 시퀀싱(Next-generation sequencing, NGS) 기술이 발전함에 따라 생성되는 데이터의 양도 계속해서 기하급수적으로 증가하고 있습니다. 이처럼 방대한 양의 데이터가 유입되면, 정확도를 떨어뜨리지 않으면서 빠르고 효율적인 분석을 실행하는 방법이 필요합니다. 대용량의 데이터와 더 향상된 딥 시퀀싱 역량을 요구하는 새로운 애플리케이션에 필요한 연산 능력의 발전을 고려하여 랩에서 바이오인포매틱스(bioinformatics, 생명정보학) 인프라를 최신 상태로 유지하는 일은 어려울 수 있습니다. Illumina는 이러한 어려움을 해소하기 위해 DRAGEN Secondary Analysis를 제공하고 있습니다. DRAGEN 소프트웨어는 랩이 이미 보유하고 있는 고성능 컴퓨팅(high-performance computing, HPC) 솔루션이나 클라우드 기반 솔루션에 손쉽게 통합되어 운영 효율성을 높여 줄 수 있습니다. 이 Technical Note는 일반적인 시나리오에서 DRAGEN 소프트웨어를 통합하는 방법을 몇 가지 예시를 통해 소개합니다.

## DRAGEN Secondary Analysis

DRAGEN 소프트웨어는 정확하고 효율적이며 포괄적인 2차 시퀀싱 데이터 분석 기능을 제공합니다. DRAGEN은 특수한 유전체(genome) 분석 알고리즘을 하드웨어 가속을 사용해 구현하는 고도로 재구성 가능한(highly reconfigurable) 필드 프로그래밍 가능 게이트 어레이(field programmable gate array, FPGA) 기술을 적용하여 만들었습니다. 지원되는 애플리케이션으로는 BCL 변환(conversion), 매핑(mapping), 정렬(alignment), 분류(sorting), 중복 리드 표시(duplicate marking), 하플로타입 변이 검출(haplotype variant calling), 유전자 정량(gene quantification), 유전자 융합(fusion) 검출, 변이 주석 달기(annotation) 등이 있습니다.

## DRAGEN 소프트웨어의 장점

- **정확한 결과** — DRAGEN은 precisionFDA Truth Challenge v2의 벤치마크 데이터를 분석했을 때 99.89%의 정확도를 달성함<sup>1</sup>
- **종합적인 솔루션** — 최대 30개의 오픈 소스 도구를 대체하는 하나의 솔루션으로 전장 유전체(whole genome), 엑솜(exome), 메틸롬(methylome), 전사체(transcriptome) 분석이 가능함
- **효율적인 분석** — 지원되는 모든 caller에서 40× 커버리지의 유전체 데이터를 약 34분 안에 처리 가능함<sup>2</sup>. DRAGEN ORA compression으로 FASTQ 파일을 최대 1/5 크기로 압축 가능함. DRAGEN Secondary Analysis는 두 가지 유전체 데이터 분석 속도 세계 기록 수립에 기여함<sup>3,4</sup>
- **비용 효율성** — 내장된 무손실 데이터 압축(lossless data compression) 기능으로 스토리지 비용을 최대 80% 줄여 주고, 사전 구성된 워크플로우는 분석 파이프라인의 개발에 드는 시간과 비용을 절감해 줌

- **멀티플랫폼 접근성** — DRAGEN Secondary Analysis는 온프레미스 서버(on-premise server), 클라우드 또는 NovaSeq™ X 시리즈, NextSeq™ 1000 및 NextSeq 2000 시스템, MiSeq™ i100 시리즈에서 실행 가능함(표 1)
- **간소화된 통합** — Illumina 시퀀싱 시스템과 쉽게 통합되므로 시퀀싱 단계부터 Illumina Connected Insights, Emedgene™, Partek™ Flow™, 또는 Correlation Engine 소프트웨어를 포함하는 3차 분석 단계까지 아우르는 간소화된 워크플로우를 지원함

## 배포 옵션

DRAGEN Secondary Analysis는 온프레미스 서버, 클라우드, 또는 기기 내 온보드 솔루션을 통해 사용할 수 있습니다. 세 가지 배포 옵션은 정확도, 속도, 또는 유연성에 미치는 영향 없이 다양한 분석 요구 사항을 충족해 줍니다.

### DRAGEN On-premise Server

DRAGEN On-premise Server는 로컬 스토리지 솔루션을 이용해 NGS 데이터를 수집하고 저장합니다(그림 1). DRAGEN Server는 특허받은 하드웨어 가속화된 유전체 분석 파이프라인이 사전 프로그래밍된 FPGA 카드를 장착한 2U 랙마운트(rackmount) 서버입니다. DRAGEN 소프트웨어의 새로운 버전은 손쉽게 다운로드 및 설치가 가능합니다. DRAGEN On-premise Server는 다음과 같은 장점을 제공합니다.

- 하나의 커맨드 라인 인터페이스(command-line interface, CLI)를 통해 DRAGEN 기능의 유연한 구성 지원
- 최대 30개의 기존 컴퓨트 인스턴스(compute instance) 대체
- 40× 커버리지의 인간 유전체 1개에 대한 NGS 데이터를 약 35분 내 처리

### HPC 시스템에 통합된 DRAGEN Server

다양한 유형의 랩에서 규모를 확장하거나 운영 효율성을 개선하기 위해 기존의 HPC 인프라에 DRAGEN Server를 통합하고 있습니다. 보유 중인 여러 컴퓨트 서버를 한번에 DRAGEN Server로 교체하든, 지원이 중단되는 서버를 개별적으로 하나씩 DRAGEN Server로 교체하든, 통합 과정은 매우 간단합니다. DRAGEN이 제공하는 가속화된 성능은 다음과 같은 간단한 4단계를 거쳐 경험할 수 있습니다.

- 데이터 센터에 DRAGEN Server를 1대 이상 추가
- DRAGEN Secondary Analysis를 사용하기 위해 분석 파이프라인 업데이트
- **[선택 사항]** DRAGEN Secondary Analysis를 포함하도록 작업 스케줄러(예: SLURM) 업데이트
- 스토리지 장치에 네트워크 파일 시스템(network file system, NFS) 및 공통 인터넷 파일 시스템(common internet file system, CIFS) 프로토콜을 지원하는 네트워크 연결 스토리지(network-attached storage, NAS) 솔루션 마운트

표 1: DRAGEN Secondary Analysis가 지원하는 광범위한 연구에 활용 가능한 2차 분석 앱<sup>a</sup>

연구용 앱	On-premise Server		Illumina 시퀀싱 시스템 내		Illumina 클라우드 플랫폼	
	DRAGEN Server	NovaSeq X 시리즈	NextSeq 1000, NextSeq 2000 시스템	MiSeq i100 시리즈	BaseSpace Sequence Hub	Illumina Connected Analytics
BCL Convert	✓	✓	✓	✓	✓	✓
DRAGEN ORA compression	✓	✓	✓	✓		곧 지원 예정
DRAGEN FASTQ + MultiQC	✓	✓	✓	✓ <sup>b</sup>	✓	✓
Whole genome	Germline + somatic	Germline + somatic	Germline + somatic		Germline + somatic	Germline + somatic
Enrichment(엑솜 포함)	Germline + somatic	Germline + somatic	Germline + somatic		Germline + somatic	Germline + somatic
DRAGEN Amplicon	✓		DNA만 지원		✓	✓
RNA	✓	✓	✓		✓	✓
Single-cell RNA	✓		✓		✓	✓
NanoString GeoMx NGS			✓		✓	
Methylation	✓	✓			✓	✓
Metagenomics	✓ <sup>c</sup>				✓	
COVID, IMAP, IMAP-FLU					✓	
TruSight Oncology 500 portfolio	✓				✓ <sup>d</sup>	✓
scATAC-Seq	✓				✓	✓
Imputation	✓				✓	✓
PGx Star Allele Caller	✓	✓	✓		✓	✓
Illumina Complete Long Reads					✓	✓
RPIP, RVEK, UPIP, VSP	✓				✓	베타
Small whole genome				✓	✓	

a. 요구되는 DRAGEN 소프트웨어 버전은 플랫폼마다 상이하므로 자세한 정보는 지역 담당자에게 문의

b. Library QC 앱을 통해 지원

c. Kmer 분류기를 기반으로 하는 메타유전체학 연구 앱으로, 앞으로 더 많은 도구 지원 예정

d. Illumina Connected Analytics 구독 필요

약어: IMAP = Illumina Microbial Amplicon Prep, RPIP = Respiratory Pathogen ID/AMR Panel, RVEK = Respiratory Virus Enrichment Kit, UPIP = Urinary Pathogen ID/AMR Panel, VSP = Viral Surveillance Panel

### 온보드

NovaSeq X 시리즈, NextSeq 1000 및 NextSeq 2000 시스템, MiSeq i100 시리즈에는 신속하고 정확한 2차 분석을 제공하는 DRAGEN 소프트웨어가 내장되어 있습니다. DRAGEN 소프트웨어는 직관적인 그래픽 사용자 인터페이스(graphical user interface, GUI)를 갖추고 있어 전문가와 비전문가가 모두 신속하게 분석을 수행하고 분석 결과를 얻을 수 있습니다. 온보드

DRAGEN 소프트웨어는 가장 널리 사용되는 앱(표 1)을 지원하는 다양한 파이프라인을 제공하며, 수상을 통해 그 성능이 입증된 머신 러닝(machine learning, ML) 및 multigenome(graph) reference 분석을 고품질 변이 검출(variant calling)을 위해 포함하고 있습니다. 온보드 DRAGEN Secondary Analysis를 사용하면 런(run)과 분석을 함께 계획할 수 있어 워크플로우가 간소화되고 터치포인트가 줄어듭니다.

확장성을 추가로 확보하려면 온보드 DRAGEN 소프트웨어를 독립형 DRAGEN Secondary Analysis 옵션(온프레미스 또는 클라우드)과 결합하여 사용해 볼 수도 있습니다. 독립형 DRAGEN 소프트웨어와 온보드 DRAGEN 소프트웨어를 결합할 경우, 확장된 워크플로우 카탈로그, 시간 절약, 처리량 극대화, 샘플 재분석 옵션 등의 장점을 누릴 수 있습니다.

**클라우드**

랩에서는 DRAGEN Server를 설치해 관리하는 대신 BaseSpace™ Sequence Hub나 Illumina Connected Analytics와 같은 Illumina의 클라우드 기반 옵션을 선택해 1차 및 2차 데이터 분석에 DRAGEN 앱을 활용해 볼 수도 있습니다(표 1). 또는 DRAGEN 소프트웨어를 Amazon Web Services(AWS), Microsoft Azure 또는 Google Cloud Platform에서 호스팅되는 랩의 프라이빗 클라우드에 호스팅하는 것도 가능합니다(그림 2).

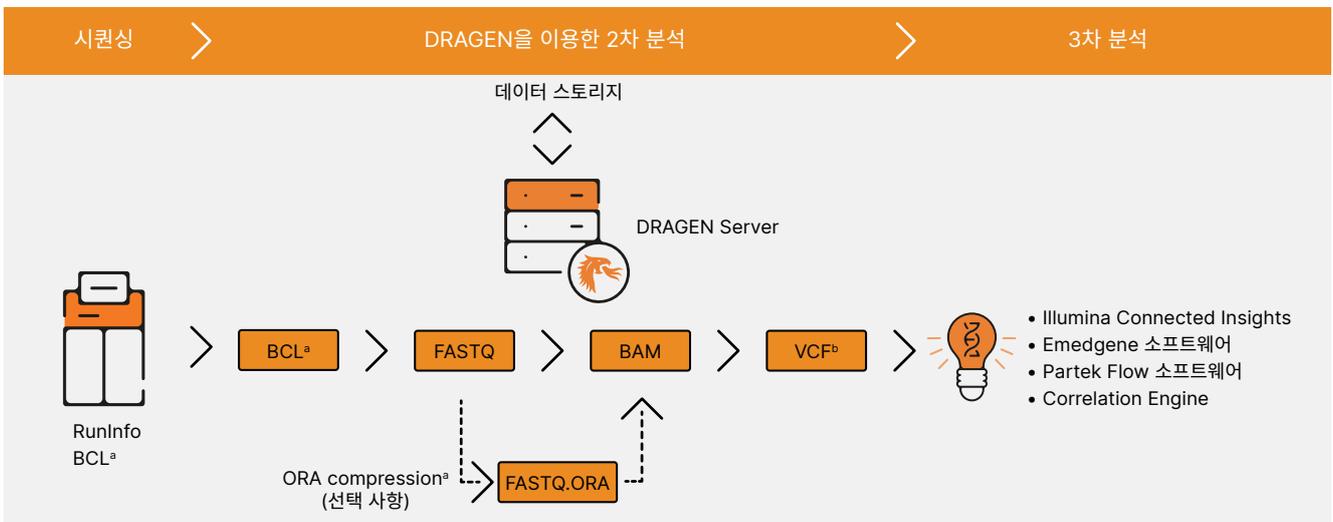


그림 1: 온프레미스에 배포된 DRAGEN Secondary Analysis — DRAGEN Server가 신속하게 데이터를 처리하고 네트워크 파일 전송을 최소화함

- a. BCL convert 및 ORA compression은 시퀀싱 시스템 내에서 온보드로 실행하거나 DRAGEN Server에서 실행 가능
- b. VCF 또는 기타 선호하는 파일 형식

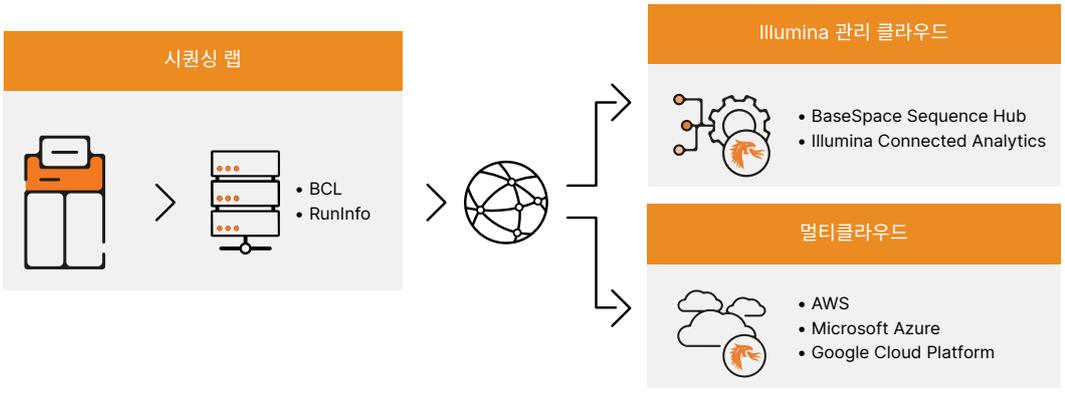


그림 2: 클라우드에 배포된 DRAGEN Secondary Analysis — BaseSpace Sequence Hub나 AWS를 통해 클라우드에서 DRAGEN 앱 사용 가능

## BaseSpace Sequence Hub

BaseSpace Sequence Hub에서 이용 가능한 클라우드 기반의 DRAGEN Secondary Analysis 앱은 정확하고 효율적인 분석뿐만 아니라 안전한 생태계와 폭넓은 확장성도 제공합니다. 랩은 규모나 연구 분야에 상관없이 BaseSpace Sequence Hub에서 DRAGEN 소프트웨어를 이용해 간편하게 버튼 조작만으로 2차 분석을 실행할 수 있습니다.

BaseSpace Sequence Hub는 Illumina 기기를 확장된 환경에서 사용할 수 있도록 해 줍니다. 기기에서 BaseSpace Sequence Hub로 암호화된 데이터가 직접 전송되므로 연구자가 큐레이션(curation)을 거친 다양한 앱을 실행해 손쉽게 데이터를 관리하고 분석할 수 있습니다. AWS 기반의 BaseSpace Sequence Hub는 다음과 같은 장점을 가지고 있습니다.

- 간편하게 버튼 조작만으로 DRAGEN Secondary Analysis를 실행하는 솔루션 포함
- 전문가와 비전문가가 모두 직관적이고 효율적으로 사용할 수 있는 GUI 적용
- 추가 인프라에 투자하지 않고도 강력한 컴퓨팅 리소스 사용 가능

## Illumina Connected Analytics

Illumina Connected Analytics는 종합적인 클라우드 기반의 바이오인포매틱스 플랫폼입니다. 연구자는 이를 이용해 안전하고 확장 가능하며 유연한 환경에서 방대한 양의 멀티오믹스(multiomics) 데이터를 관리, 분석 및 해석할 수 있습니다. 연구자는 Illumina Connected Analytics에서 사전 패키지(prepackaged)된 파이프라인 또는 커스텀 파이프라인에 통합할 개별 도구로 DRAGEN Secondary Analysis 제품군을 사용할 수 있습니다.

## 멀티클라우드

FPGA가 지원되는 리전(region)에서는 멀티클라우드에서 DRAGEN을 사용할 수 있습니다. 인스턴스 구성 방법은 [DRAGEN Multi-Cloud User Guide](#)를 참조하시기 바랍니다.

## 요약

랩은 기존의 분석 인프라에 DRAGEN Secondary Analysis를 통합함으로써 속도, 정확도, 효율성, 호환성, 확장성을 모두 확보할 수 있습니다. 다양한 통합 시스템의 구현을 위해 제공되는 여러 옵션 중 하나를 선택하여 가능한 경우 데이터 수용력을 높여 볼 수도 있습니다. Illumina는 랩에서 요구 사항을 충족하는 솔루션을 선택하는 방법을 안내해 드릴 수 있습니다.

## 상세 정보

### DRAGEN Secondary Analysis

## 참고 문헌

1. Illumina. DRAGEN sets new standard for data accuracy in PrecisionFDA benchmark data. Optimizing variant calling performance with Illumina machine learning and DRAGEN graph. [illumina.com/science/genomics-research/articles/dragen-shines-again-precisionfda-truth-challenge-v2.html](https://illumina.com/science/genomics-research/articles/dragen-shines-again-precisionfda-truth-challenge-v2.html). Published January 12, 2022. Accessed October 28, 2024.
2. Data on file. Illumina, Inc. 2024.
3. BioIT World. Children's Hospital of Philadelphia, Edico set world record for secondary analysis speed. [bio-itworld.com/news/2017/10/23/children-s-hospital-of-philadelphia-edico-set-world-record-for-secondary-analysis-speed](https://bio-itworld.com/news/2017/10/23/children-s-hospital-of-philadelphia-edico-set-world-record-for-secondary-analysis-speed). Published October 23, 2017. Accessed October 28, 2024.
4. San Diego Union Tribune. Rady Children's Institute sets Guinness world record. [sandiegouniontribune.com/2018/02/12/rady-childrens-institute-sets-guinness-world-record/](https://sandiegouniontribune.com/2018/02/12/rady-childrens-institute-sets-guinness-world-record/). Published February 12, 2018. Accessed October 28, 2024.



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

© 2025 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-00316 KOR