

# Illumina Complete Long Read Prep, Human

롱 리드 인간 전장 유전체  
시퀀싱을 위한  
확장 가능한  
고성능 솔루션

- 한 대의 기기에서 롱 리드와 쇼트 리드의 포괄적인 인간 전장 유전체 시퀀싱 수행 가능
- 최적화된 라이브러리 준비 및 데이터 분석 성능을 바탕으로 매우 정확하고 신뢰할 수 있는 결과 도출
- 적은 양의 DNA를 사용하여 하루 안에 완료되는 자동화 지원 워크플로우

illumina®

## 소개

차세대 시퀀싱(Next-generation sequencing, NGS)은 과학자에게 유전체(genome)를 해독하여 생명 활동에 대한 이해의 폭을 넓힐 수 있는 기회를 제공합니다. Illumina의 입증된 sequencing by synthesis(SBS) chemistry와 수상 경력에 빛나는 DRAGEN™ 2차 분석 기능이 만나 정확도가 매우 높은 전장 유전체 시퀀싱(whole-genome sequencing, WGS) 데이터를 생성합니다.<sup>1</sup> 최근 DRAGEN 분석 기술이 발전함에 따라 가장 매핑하기 어려운(difficult-to-map) 유전자에 남아 있던 갭이 점점 좁혀지고 있습니다.<sup>2,3</sup> 그중에서도 극히 일부의 분석이 어려운 유전자 영역에는 높은 정확도와 해상도로 매핑이 가능한 긴 리드 길이(read length)를 적용하면 한결 더 좋은 결과를 기대할 수 있습니다. 그러나 다수의 롱 리드 시퀀싱(long-read sequencing) 솔루션은 많은 양의 DNA 사용을 요구하고, 처리량이 낮은(low-throughput) 복잡한 워크플로우를 지원하며, 결과의 변동성이 매우 심하다는 단점이 있어 활용이나 도입이 쉽지 않았습니다.<sup>4-7</sup>

Illumina Complete Long Reads 제품은 유전체 검사실에서 롱 리드 시퀀싱을 간결하고 손쉽게 적용할 수 있도록 해 줍니다. Illumina Complete Long Prep, Human은 이러한 새로운 롱 리드 기술을 기반으로 한 최초의 제품입니다. 이 고성능 WGS assay는 일반적인 NGS 워크플로우를 사용하여 NovaSeq™ 6000 시스템이나 NovaSeq X 시리즈에서 contiguous 롱 리드 시퀀스를 생성합니다(그림 1). 하루 안에 완료 가능한 효율적인 라이브러리 준비(library prep) 프로토콜은 높은 처리량(high-throughput)을 요구하는 연구에 맞춰 손쉽게 확장이 가능하며 50 ng의 DNA만을 사용하고 전문적인 추출(extraction), 절단(shearing), 크기 선별(size selection) 과정이 없습니다.

## NovaSeq 플랫폼으로 생성하는 고품질의 롱 리드

Illumina Complete Long Reads Prep, Human은 NovaSeq X Plus, NovaSeq X 및 NovaSeq 6000 시퀀싱 시스템에서 사용이 가능합니다. 따라서 사용자는 한 대의 기기로 롱 리드 데이터와 쇼트 리드 데이터를 모두 생성할 수 있습니다(표 1). 이 유연한 assay는 다양한 품질의 샘플에서 일관된 결과를 제공하며 다른 롱 리드 솔루션보다 90% 적은 DNA 사용량을 요구합니다(그림 2). 또한 assay가 일반적인 억제제(inhibitor)나 오염 물질에 대한 저항성이 있기 때문에 혈액, 타액 또는 조직에서 얻은 DNA에도 적합합니다.

Illumina Complete Long Reads는 Illumina의 사유 library prep assay, 입증된 SBS chemistry 그리고 강력한 DRAGEN 2차 분석을 결합하여 매우 정확한 롱 리드 데이터를 생성합니다(그림 3). 긴 단일 분자(single-molecule) DNA 절편은 효소를 이용하여 "land-mark"라고 불리는 고유한 패턴으로 마크된 후 증폭(amplification) 및 시퀀싱 과정을 거칩니다. 소프트웨어는 land-mark를 기준으로 반복 영역이나 매핑이 어려운 영역을 구별하고 N50이 5~7 kb인 롱 리드를 생성합니다(그림 2). 롱 리드 데이터는 unmarked WGS 라이브러리와 결합되어 원래의 단일 분자 절편을 완전히 정확하게 표현하는 길이가 긴 contiguous 리드가 만들어집니다.



그림 1: Illumina Complete Long Read 워크플로우 — 확장 가능하고 최적화된 라이브러리 준비 프로토콜을 통해 NovaSeq 6000 시스템 또는 NovaSeq X 시리즈로 시퀀싱 후 DRAGEN 2차 분석을 수행하여 포괄적인 롱 리드 WGS 데이터 생성. Illumina Complete Long Read 라이브러리는 표준 30× WGS 라이브러리의 준비, 시퀀싱 및 분석이 반드시 동반되어야 함.

a. 최소 10 ng으로 사용 가능하나, 50 ng의 DNA 사용이 권장됨.

표 1: Illumina Complete Long Read assay 권장 샘플 처리량<sup>a,b,c</sup>

300-Cycle Reagent Kit	플로우 셀당 샘플 수	플로우 셀당 데이터 아웃풋	런 타임
NovaSeq 6000 S4 Reagent	4개	약 3 Tb	약 44시간
NovaSeq X Series 10B Reagent	4개	약 3 Tb	약 24시간
NovaSeq X Series 25B Reagent	10~11개	약 8 Tb	약 48시간

- a. 2 × 150 bp 시퀀싱 런(run) 조건에서 대략적으로 Illumina Complete Long Reads의 30× 커버리지(coverage) 생성.
- b. 분석 시 동일한 샘플로 얻은 표준 30× 쇼트 리드 인간 전장 유전체 데이터 필요. [Illumina DNA PCR-Free Prep](#) 사용 권장. 타사 WGS 키트 호환 가능. Unmarked 라이브러리의 준비 또는 개별 시퀀싱 불필요. 이전 런 샘플에서 생성된 FASTQ 파일 사용 가능.
- c. Illumina Complete Long Read 라이브러리를 NovaSeq 플랫폼에서 시퀀싱할 경우 런의 Q30 점수가 NovaSeq 사용보다 낮게 보고될 수 있으나, 이는 시퀀싱 런이나 라이브러리의 성능 문제를 의미하지 않음.

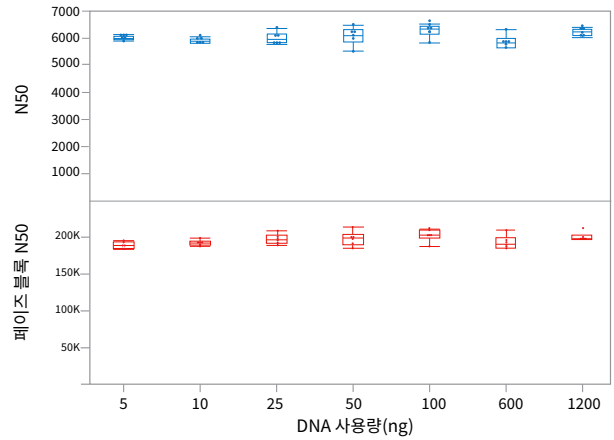


그림 2: DNA 사용량 관계없이 일관된 결과를 제공하는 Illumina Complete Long Read assay — 5~1200 ng의 DNA 사용 시 생성되는 N50과 페이즈 블록(phase block) N50의 데이터 품질이 비슷한 것을 확인 가능. N50이란 전체 어셈블리(assembly) 길이의 50%일 때 가장 짧은 콘티그(contig)의 시퀀스 길이로 정의됨. N50은 페이징 블록의 크기를 측정하는 척도로도 사용 가능.

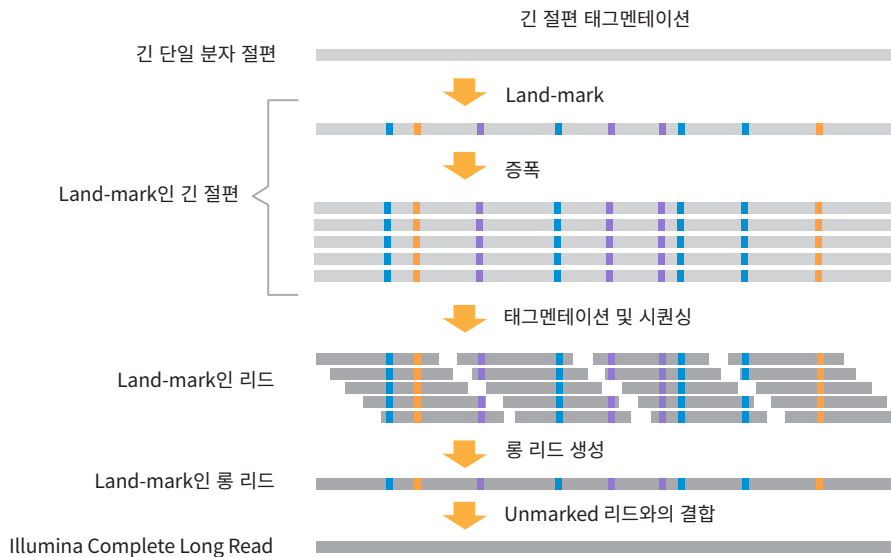


그림 3: Illumina Complete Long Read assay의 원리 — 태그멘테이션(Tagmentation)<sup>8</sup>을 통해 긴 DNA 절편(> 10 kb)을 생성하므로 별도의 절단이나 크기 선별 과정이 필요하지 않음. 긴 절편은 단일 분자 수준에서 land-mark가 되어 바코드나 어댑터(adapter) 없이도 절편 내 롱 리드 정보를 포착 및 보존함. Land-mark인 긴 절편은 증폭 과정을 거치며, 2차 태그멘테이션 단계를 통해 시퀀싱 가능한 라이브러리가 준비됨. 데이터 분석 단계에서는 강력한 DRAGEN 소프트웨어가 롱 리드 데이터를 생성한 후 이를 표준 unmarked WGS 라이브러리(동일한 샘플을 개별적으로 시퀀싱한 경우)와 결합하여 정확도가 매우 높은 완전한 롱 리드를 제공함.

### 정확도 높은 WGS 데이터 확보

Illumina Complete Long Read 데이터는 일반적인 쇼트 리드 WGS 데이터를 보완하며, 다음을 통해 한층 더 포괄적인 전장 유전체 데이터를 제공합니다.

- 상동성(Homology)이 높거나 반복되는 영역이 있는 매핑이 어려운 영역에서의 변이 검출
- 복잡한 구조적 변이(structural variant), 유사 유전자(pseudogene), 큰 삽입/결실(insertion/deletion, Indel)에 유용
- 변이 페이징(phasing), 하플로타입(haplotype) 검출

Illumina DNA PCR-Free Prep을 사용해 준비한 라이브러리를 DRAGEN 2차 분석의 정확도 높은 인간 WGS 데이터와 비교했을 때, Illumina Complete Long Reads 데이터는 다양한 변이 유형에 대해 향상된 변이 검출 정확도를 보입니다<sup>3</sup>(그림 4). PrecisionFDA Truth Challenge v2 데이터 세트를 살펴보면, Illumina Complete Long Reads assay 사용 시 WGS 정밀도(precision) 및 재현율(recall)을 나타내는 F1 점수는 99.87%였습니다(그림 5).<sup>9,10</sup>

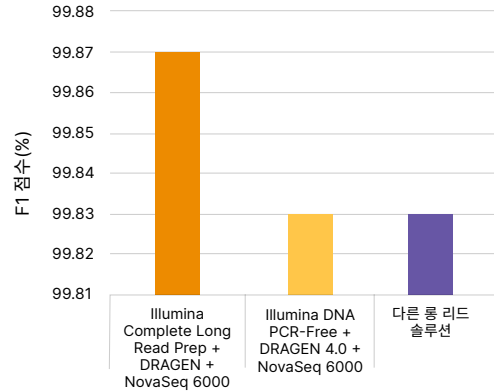


그림 5: 정확도의 새로운 기준 — WGS에 대한 정밀도와 재현율을 반영하는 F1 점수(%)를 기준으로 볼 때, Illumina Complete Long Read Prep, Human은 훨씬 더 우수한 변이 검출 정확도를 보임.

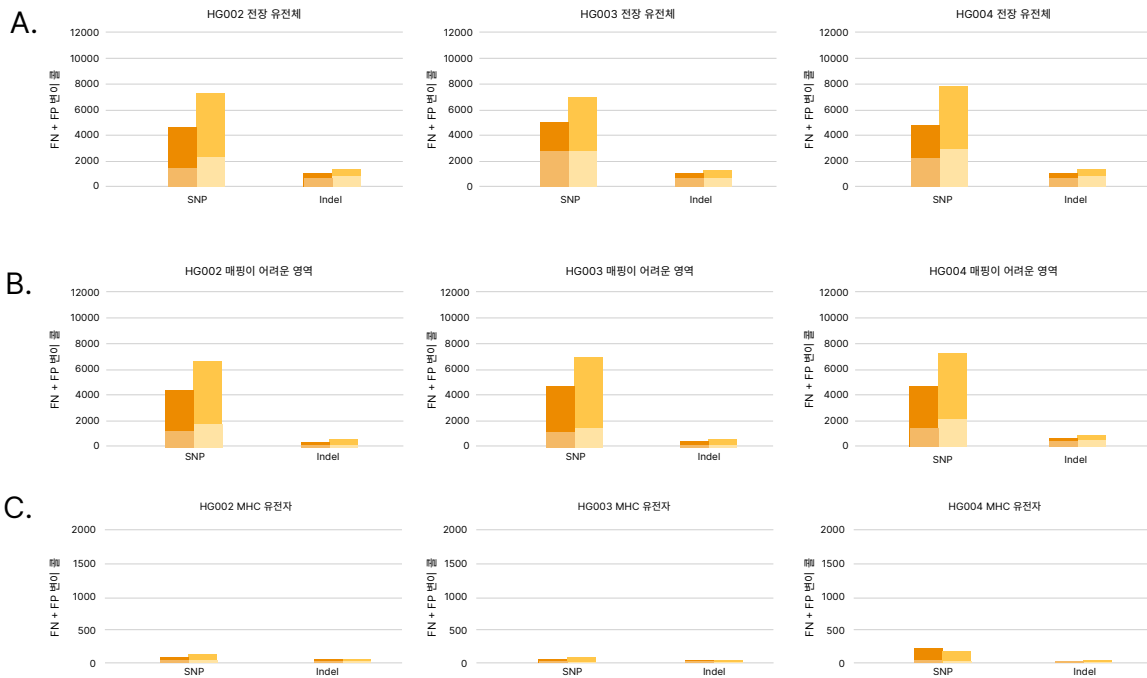


그림 4: 분석이 어려운 유전자 영역에서 매우 정확한 변이 검출 역량을 보이는 Illumina Complete Long Reads assay — Genome in a Bottle 컨소시엄의 HG002, HG003 및 HG004 인간 레퍼런스 샘플에 대한 위양성(false positive, FP) 및 위음성(false negative, FN) 결과로 측정된 단일 염기 다형성(single nucleotide polymorphism, SNP) 및 Indel 변이 검출 정확도.<sup>11</sup> (A) 전장 유전체, (B) 매핑이 어려운 영역, (C) MHC 유전자에 대한 Illumina Complete Long Reads assay(주황색)와 Illumina DNA PCR-Free Prep(노란색)의 WGS 데이터를 비교한 그래프.

## 높은 처리량을 요구하는 연구에 적합한 확장 가능하고 간소화된 워크플로우

ILLUMINA Complete Long Read Prep 워크플로우는 더 많은 샘플을 사용하는 포괄적인 WGS를 지원하기 위해 손쉽게 확장이 가능하도록 설계되었습니다. 라이브러리 준비 프로토콜이 간단하여 약 8시간(수작업 시간 약 4시간) 안에 완료 가능하며 일반적인 검사실 장비만을 요구하고 자동화가 쉽습니다. NovaSeq X Plus 시스템 사용 시 연간 최대 3천 개의 정확도 높은 유전체를 생성할 수 있습니다.\* ILLUMINA Complete Long Reads는 기존의 WGS 데이터 세트를 보강하거나 변이에 대한 심층적인 연구가 필요할 때 추가 도구로 활용할 수 있습니다.

## 요약

롱 리드 데이터는 유전체에서 가장 분석이 어려운 영역에 유용한 정보를 제공할 수 있습니다. ILLUMINA Complete Long Read Prep, Human은 유전체 검사실에서 한 대의 기기로 롱 리드와 쇼트 리드를 모두 분석할 수 있도록 해주어 포괄적인 WGS에 대한 접근성을 높여줍니다. 간소화되고 익숙한 워크플로우와 ILLUMINA의 입증된 SBS chemistry 및 DRAGEN 분석 기술은 함께 시너지 효과를 발휘해 우수한 확장성과 정확성을 갖춘 전장 유전체 시퀀싱을 제공합니다. ILLUMINA는 앞으로 더 많은 제품의 출시를 통해 ILLUMINA Complete Long Reads의 유용성을 높일 예정입니다.

## 상세 정보

[ILLUMINA Complete Long Read Prep, Human](#)

[롱 리드 시퀀싱 기술](#)

## 참고 문헌

- Mehio R, Ruehle M, Catreux S, et al. DRAGEN Wins at Precision- FDA Truth Challenge V2 Showcase Accuracy Gains from Alt-aware Mapping and Graph Reference Genomes. [illumina.com/science/genomics-research/dragen-wins-precisionfda- challenge-showcase-accuracy-gains.html](https://www.illumina.com/science/genomics-research/dragen-wins-precisionfda- challenge-showcase-accuracy-gains.html). Accessed January 12, 2023.
- ILLUMINA. DRAGEN Bio-IT Platform data sheet. [illumina.com/content/dam/illumina/gcs/assembled-assets/marketing-literature/dragen-bio-it-data-sheet-m-gl-00680/dragen-bio-it-data-sheet-m-gl-00680.pdf](https://www.illumina.com/content/dam/illumina/gcs/assembled-assets/marketing-literature/dragen-bio-it-data-sheet-m-gl-00680/dragen-bio-it-data-sheet-m-gl-00680.pdf). Published 2022. Accessed February 1, 2023.
- ILLUMINA. Accuracy improvements in germline small variant calling with the DRAGEN Bio-IT Platform. [illumina.com/content/dam/illumina/gcs/assembled-assets/marketing-literature/dragen-v4-accuracy-app-note-m-gl-01016/dragen-v4-accuracy-app-note-m-gl-01016.pdf](https://www.illumina.com/content/dam/illumina/gcs/assembled-assets/marketing-literature/dragen-v4-accuracy-app-note-m-gl-01016/dragen-v4-accuracy-app-note-m-gl-01016.pdf). Published 2022. Accessed February 1, 2023.
- Pacific Biosciences. Preparing DNA for PacBio HiFi sequencing—Extraction and quality control. [pacb.com/wp-content/uploads/Technical-Note-Preparing-DNA-for-PacBio-HiFi-Sequencing-Extraction-and-Quality-Control.pdf](https://www.pacb.com/wp-content/uploads/Technical-Note-Preparing-DNA-for-PacBio-HiFi-Sequencing-Extraction-and-Quality-Control.pdf). Published 2022. Accessed January 12, 2023.
- Pacific Biosciences. Preparing whole genome and metagenome libraries using SMRTbell prep kit 3.0. [pacb.com/wp-content/uploads/Procedure-checklist-Preparing-whole-genome-and-metagenome-libraries-using-SMRTbell-prep-kit-3.0.pdf](https://www.pacb.com/wp-content/uploads/Procedure-checklist-Preparing-whole-genome-and-metagenome-libraries-using-SMRTbell-prep-kit-3.0.pdf). Published 2022. Accessed January 12, 2023.
- Oxford Nanopore Technologies. Ligation Sequencing Kit. [store.nanoporetech.com/us/ligation-sequencing-kit110.html](https://www.store.nanoporetech.com/us/ligation-sequencing-kit110.html). Accessed January 12, 2023.
- Pacific Biosciences. Low Yield Troubleshooting Guide. [pacb.com/wp-content/uploads/Guide-Low-Yield-Troubleshooting.pdf](https://www.pacb.com/wp-content/uploads/Guide-Low-Yield-Troubleshooting.pdf). Published 2018. Accessed January 12, 2023.
- ILLUMINA. Tagmentation. [illumina.com/techniques/sequencing/ngs-library-prep/tagmentation.html](https://www.illumina.com/techniques/sequencing/ngs-library-prep/tagmentation.html). Accessed January 12, 2023.
- ILLUMINA. Data on file. 2022.
- PrecisionFDA. Truth Challenge V2: Calling Variants from Short and Long Reads in Difficult-to-Map Regions. [precision.fda.gov/challenges/10](https://www.precision.fda.gov/challenges/10). Accessed January 12, 2023.
- Genome in a Bottle. National Institute of Standards and Technology. [nist.gov/programs-projects/genome-bottle](https://www.nist.gov/programs-projects/genome-bottle). Accessed January 12, 2023.

\* ILLUMINA Complete Long Read Prep, Human을 NovaSeq X Plus 시스템 및 25B 플로우 셀과 함께 사용해 듀얼 플로우 셀 런을 수행한 경우 달성 가능한 처리량

제품 목록

제품	카탈로그 번호
Illumina Complete Long Read Prep, Human (24 samples)	20089108
Illumina Complete Long Read Prep, Human (8 samples)	20086823
Illumina Unique Dual Indexes, LT (48 indexes, 48 samples)	20098166
Illumina Complete Long Read WGS App	20100421

illumina®

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

© 2023 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-GL-01420 v1.1 KOR