DATA SHEET

Illumina Complete Long Read Prep, Human

A high-performance, scalable solution for long-read human whole-genome sequencing

- Comprehensive human WGS with both long and short reads from the same instrument
- Optimized library prep and analysis performance for highly accurate, reliable results
- One-day, automation-compatible workflow with low DNA input requirements

illumina®

Introduction

Next-generation sequencing (NGS) enables scientists to decipher the genome for a deeper understanding of biology. Proven Illumina sequencing by synthesis (SBS) chemistry combined with award-winning DRAGEN™ secondary analysis delivers whole-genome sequencing (WGS) data with outstanding accuracy.[1] The latest advancements in DRAGEN analysis are closing remaining gaps across the most difficult-to-map genes.[2,3] Still, a very small fraction of genic regions can benefit from longer read lengths to enable accurate resolution and mapping of these challenging regions. However, many long-read sequencing solutions have been plagued by high DNA input requirements, complex workflows with low throughput, and highly variable results, and these have limited their utility and adoption.[4-7]

Illumina Complete Long Reads make long-read sequencing accessible and streamlined for genomic labs. Illumina Complete Long Prep, Human is the first product based on this novel long-read technology. The high-performance WGS assay uses a standard NGS workflow to generate contiguous long-read sequences on the NovaSeq™ 6000 System and NovaSeq X Series (Figure 1). The efficient, single-day library preparation protocol is easy to scale for high-throughput studies and requires only 50 ng DNA input with no specialized extractions, shearing, or size selection.

Generate high-quality long reads on NovaSeq platforms

Illumina Complete Long Read Prep, Human is compatible with the NovaSeq X Plus, NovaSeq X, and NovaSeq 6000 Sequencing Systems, giving users access to both long- and short-read data on the same instrument (Table 1). The flexible assay delivers consistent results across samples of variable quality, while requiring 90% less DNA input than other long-read solutions (Figure 2). Because the assay is resistant to common inhibitors and contaminants, it works well with DNA from blood, saliva, or tissue.

Illumina Complete Long Reads combines a proprietary library prep assay, proven Illumina SBS chemistry, and powerful DRAGEN secondary analysis to generate highly accurate long-read data (Figure 3). Long, single-molecule DNA fragments are enzymatically marked with unique patterns (or "land-marks"), then amplified and sequenced. The land-marks allow software to distinguish repetitive or difficult-to-map regions and generate long reads with an N50 of 5–7 kb (Figure 2). Long-read data are combined with an unmarked WGS library to produce long contiguous reads that are a complete and accurate representation of the original single-molecule fragment.

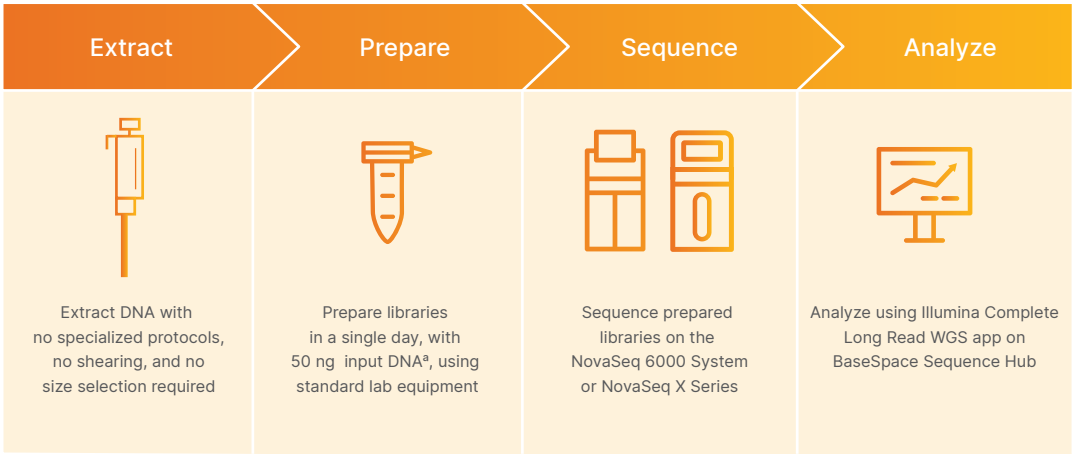

Figure 1: Illumina Complete Long Read workflow—Access comprehensive long-read WGS data using a scalable, optimized library prep protocol, sequencing on the NovaSeq 6000 System or NovaSeq X Series, and DRAGEN secondary analysis. A standard 30× WGS library should be prepared, sequenced, and analyzed with the Illumina Complete Long Read library.

a. While inputs as low as 10 ng are possible, Illumina recommends 50 ng of DNA

Table 1: Recommended sample throughput for Illumina Complete Long Read assay[a,b,c]

300-cycle reagent kits	Samples per flow cell	Output per flow cell	Run time
NovaSeq 6000 S4 Reagents	4	~3 Tb	~44 hr
NovaSeq X Series 10B Reagents	4	~3 Tb	~24 hr
NovaSeq X Series 25B Reagents	10–11	~8 Tb	~48 hr

a. Requires 2 × 150 bp sequencing run, generating approximately 30× coverage of Illumina Complete Long Reads.
b. Requires 30× standard short-read human whole-genome data from the same sample for analysis. Illumina DNA PCR-Free Prep is recommended. Third-party WGS kits are also compatible. Unmarked library does not need to be prepared or sequenced in parallel; can use FASTQ files from a previously run sample.
c. Sequencing Illumina Complete Long Read libraries on NovaSeq platforms may cause the reported Q30 score of a run to fall below the NovaSeq specification. This does not indicate a performance issue with the sequencing run, nor the library.

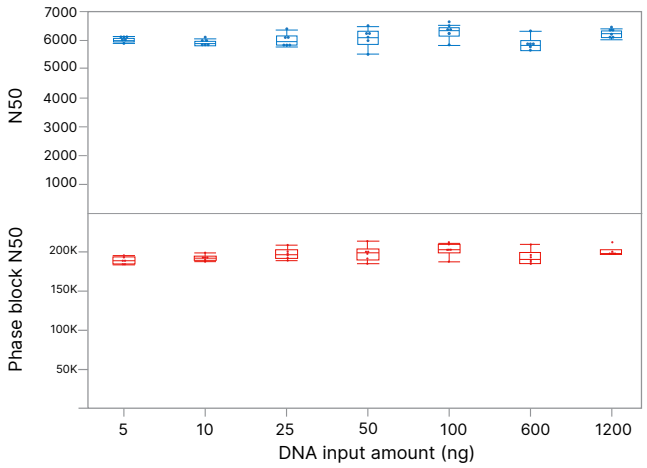

Figure 2: Illumina Complete Long Read assay delivers consistent results across DNA input amounts—DNA inputs from 5 ng to 1200 ng generate similar data quality for N50 and phase block N50. N50 is defined as the sequence length of the shortest contig at 50% of the total assembly length. It can also be used as a measure of the size of phasing blocks.

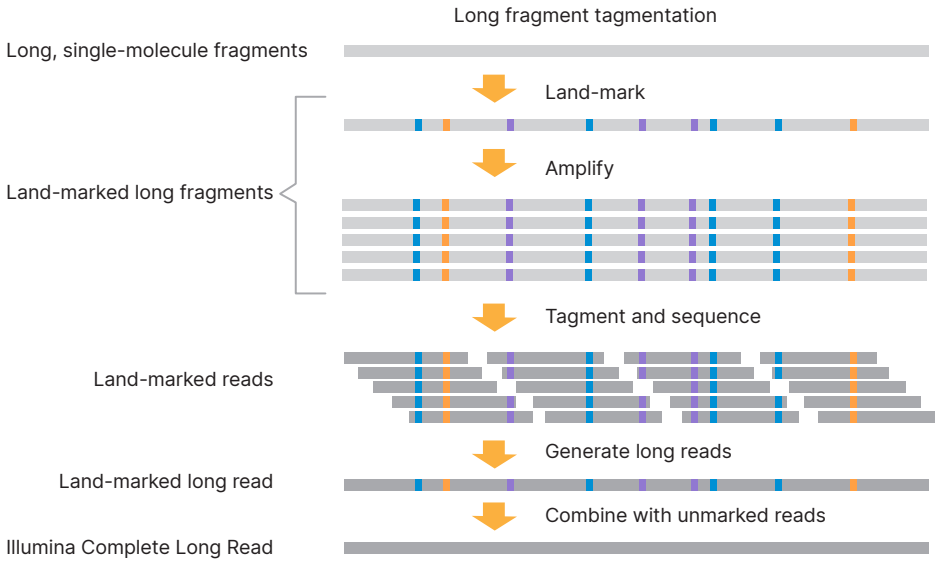

Figure 3: How the Illumina Complete Long Reads assay works—The assay uses tagmentation[8] to make long DNA fragments (> 10 kb), eliminating the need for shearing or size selection. Long fragments are "land-marked" at the single-molecule scale to capture and preserve long-read information within the fragment (without complex barcodes or adapters). Land-marked long fragments are amplified, then a second tagmentation step prepares the libraries for sequencing. During analysis, powerful DRAGEN software generates long reads and combines the data with a standard, unmarked WGS library (from the same sample, sequenced separately) to produce highly accurate complete long reads.

Access highly accurate WGS

Illumina Complete Long Read data complements standard short-read WGS data and delivers more comprehensive whole genomes by:

- Calling variants in challenging, difficult-to-map regions with high homology or repetitive regions
- Resolving complex structural variants, pseudogenes, and large insertion–deletions (indels)
- Phasing variants and calling haplotypes

Illumina Complete Long Read data demonstrate improved variant calling accuracy across variant types compared to high-accuracy human WGS data generated using Illumina DNA PCR-Free Prep for library preparation and DRAGEN secondary analysis[3] (Figure 4). With PrecisionFDA Truth Challenge v2 data sets, the F1 score reflecting precision and recall for WGS using the Illumina Complete Long Read assay was 99.87% (Figure 5).[9,10]

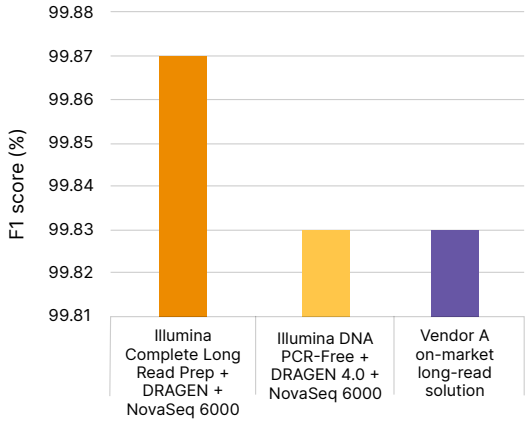

Figure 5: A new standard for accuracy—Illumina Complete Long Read Prep, Human delivers unprecedented accuracy for variant calling, as measured by F1 score (%), reflecting precision and recall for WGS.

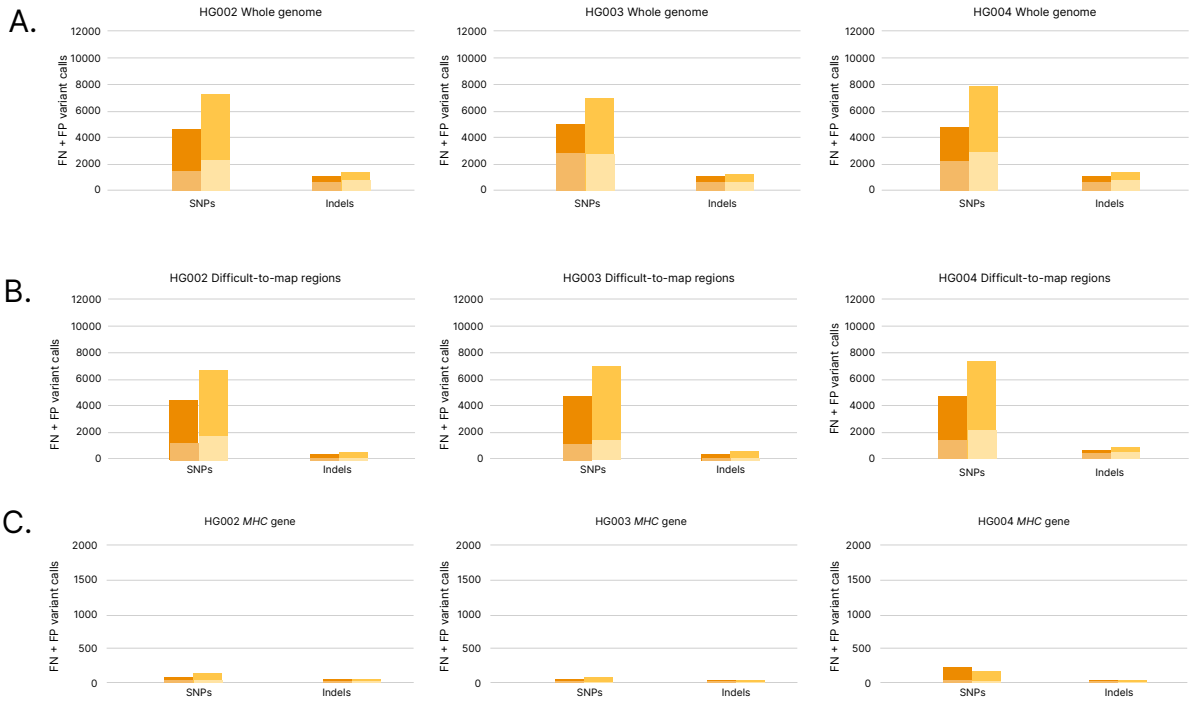

Figure 4: Illumina Complete Long Read assay performs highly accurate variant calling for challenging genic regions—Single nucleotide polymorphisms (SNP) and indel variant calling accuracy measured as false positives (FP) and false negatives (FN) for Genome in a Bottle human reference samples HG002, HG003, and HG004.[11] Comparing WGS data from Illumina Complete Long Read assay (orange) and Illumina DNA PCR-Free Prep (yellow) across (A) whole genome, (B) difficult-to-map regions, and (C) *MHC* gene.

Scalable, streamlined workflow for high-throughput studies

The Illumina Complete Long Read Prep workflow is highly scalable to support comprehensive WGS for more samples. The simple library prep protocol takes approximately 8 hours (with ~ 4 hours hands-on time), requires only standard lab equipment, and is easy to automate. With the NovaSeq X Plus System, users can generate up to 3000 high-accuracy genomes per year.* Illumina Complete Long Reads can be used to augment existing WGS data sets or as a reflex tool for deeper variant discovery.

Summary

Long-read information can help resolve the most challenging regions of the genome. Illumina Complete Long Read Prep, Human makes comprehensive WGS easily accessible for genomics labs by enabling both long- and short-reads on the same instrument. The streamlined, familiar workflow and synergy with proven Illumina SBS chemistry and DRAGEN analysis deliver the most scalable and accurate whole genome on the market. Future product offerings will expand the utility of Illumina Complete Long Reads.

Learn more

Illumina Complete Long Read Prep, Human

Long-read sequencing technology

* Potential throughput when Illumina Complete Long Read Prep, Human is used with the NovaSeq X Plus System, dual flow cell run with 25B flow cells.

References

1. 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. Accessed January 12, 2023.
2. 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. Published 2022. Accessed February 1, 2023.
3. 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. Published 2022. Accessed February 1, 2023.
4. 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. Published 2022. Accessed January 12, 2023.
5. 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. Published 2022. Accessed January 12, 2023.
6. Oxford Nanopore Technologies. Ligation Sequencing Kit. store.nanoporetech.com/us/ligation-sequencing-kit110.html. Accessed January 12, 2023.
7. Pacific Biosciences. Low Yield Troubleshooting Guide. pacb.com/wp-content/uploads/Guide-Low-Yield-Troubleshooting.pdf. Published 2018. Accessed January 12, 2023.
8. Illumina. Tagmentation. illumina.com/techniques/sequencing/ngs-library-prep/tagmentation.html. Accessed January 12, 2023.
9. Illumina. Data on file. 2022.
10. PrecisionFDA. Truth Challenge V2: Calling Variants from Short and Long Reads in Difficult-to-Map Regions. precision.fda.gov/challenges/10. Accessed January 12, 2023.
11. Genome in a Bottle. National Institute of Standards and Technology. nist.gov/programs-projects/genome-bottle. Accessed January 12, 2023.

Ordering information

Product	Catalog no.
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®

1.800.809.4566 toll-free (US) | +1.858.202.4566 tel
techsupport@illumina.com | www.illumina.com

M-GL-01420 v1.1