DATA SHEET

Illumina Complete Long Read Prep with Enrichment, Human

Flexible, cost-effective, and accurate long reads on your Illumina sequencing system

- Targeted long reads coupled with short reads for complementary, enhanced human genome insights
- Scalable, robust library prep with reliable results from low DNA input
- Automation-compatible two-day library prep workflow without the need for specialized equipment

illumina®

Introduction

Proven Illumina sequencing by synthesis (SBS) chemistry combined with award-winning DRAGEN™ secondary analysis delivers whole-genome sequencing (WGS) data with outstanding accuracy.[1,2] Still, small portions of the genome, including highly homologous or repetitive regions, are difficult to map with short reads alone. In these cases, long-read sequencing can complement standard short-read WGS data. Using high-accuracy short reads and long reads together can help deliver enhanced resolution and mapping of historically challenging regions.

In the past, long-read sequencing has been incompatible with many sample types due to strict DNA quality and high input requirements.[3-6] Illumina Complete Long Reads technology now makes long-read sequencing accessible for genomics labs by enabling both short reads and complementary long reads on the same system with a single analysis pipeline. This high-performance assay uses a standard Illumina next-generation sequencing (NGS) workflow to generate contiguous long-read sequences and requires only 10 ng DNA input with no specialized extractions, shearing, or size selection (Figure 1).

Illumina Complete Long Read Prep with Enrichment, Human creates a cost-effective, flexible solution for targeted long-read sequencing. The efficient, two-day library preparation with enrichment is automation-friendly and easy to scale for high-throughput studies.

Targeted long reads can be used to address regions known to have low mappability with standard short-read sequencing. Alternatively, targeted long reads can be applied across entire genes or regions to enable phased sequencing of up to hundreds of kilobases to resolve haplotypes. Use Illumina Complete Long Read Prep with Enrichment, Human to augment existing WGS data sets as a reflex tool for broader variant detection.

Targeted, high-quality long reads from your NovaSeq™ platform

Illumina Complete Long Read Prep with Enrichment, Human is validated with the NovaSeq X Plus, NovaSeq X, NovaSeq 6000, and NovaSeq 6000Dx (RUO mode) sequencing systems. The flexible assay delivers consistent results across samples of variable quality, while requiring 90% less DNA input than other long-read sequencing solutions. Illumina recommends 50 ng of input DNA, yet robust results are possible with as little as 10 ng. Illumina Complete Long Read technology is resistant to common inhibitors and contaminants and works well with DNA from blood, saliva, or tissue.[7,8] This enables comprehensive insights from more sample types than with other long-read solutions.

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 2). Long, single-molecule

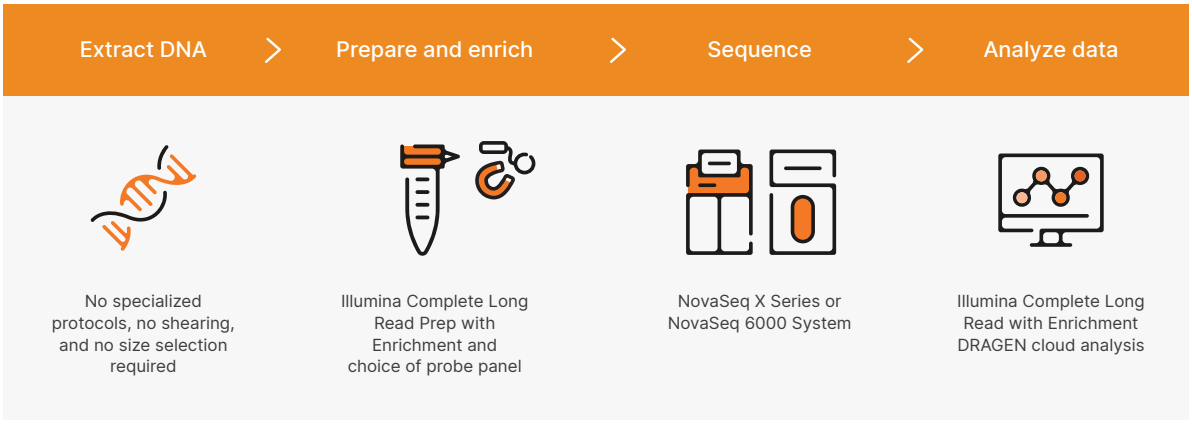


Figure 1: Illumina Complete Long Reads with Enrichment workflow—Access cost-effective, targeted long-read WGS data using a scalable, optimized library prep protocol, proven Illumina sequencing chemistry, and DRAGEN secondary analysis. Requires ≥ 30× standard short-read WGS data from the same sample for analysis. FASTQ files from a previously run sample can be used.

DNA fragments are marked with unique patterns (or "land-marks"), then amplified, enriched, and sequenced. Adding a hybrid–capture enrichment step with targeted probe panels allows sequencing to focus on regions that benefit most from long-read information. The land-marks enable differentiation of repetitive or difficult-to-map regions to generate long reads with an N50 of 5–7 kb.[7] Long-read data are combined with a standard unmarked WGS library to produce long contiguous reads that are a complete and accurate representation of the targeted original single-molecule fragments.

Long reads where you need them

Targeted long reads layered onto high-accuracy short-read WGS data help researchers focus sequencing dollars on resolving known challenging regions of the genome. Targeted long reads offer high flexibility with choice of multiple predesigned panels or custom panels enabled by Illumina DesignStudio™ software, a free, user-friendly assay design tool (Table 1, Table 2). Choose the enrichment probe panel that best meets your research needs: either to enhance regions of low coverage or phase regions of the genome.

Illumina Human Comprehensive Panel

The large Illumina Human Comprehensive Panel addresses the small portion of genic regions that would benefit from long reads, targeting low-coverage spots across > 6500 protein-coding genes.[9] In developing this panel, Illumina evaluated the full set of over 20,000 protein-coding genes to target regions of low mappability. Genes that are comprehensively covered by short reads alone were excluded. The Human Comprehensive Panel delivers enhanced coverage and variant calling abilities across target regions (Figure 3A). This panel is optimized for use with Illumina Complete Long Read Prep with Enrichment, Human and available premanufactured or as a predesigned panel that can be customized in the DesignStudio tool.

Predesigned panels with DesignStudio software

For a lower cost, higher throughput option, researchers can choose focused panels for long-read enrichment (Table 1). Several predesigned panels are available in DesignStudio software and can be customized (Table 2).

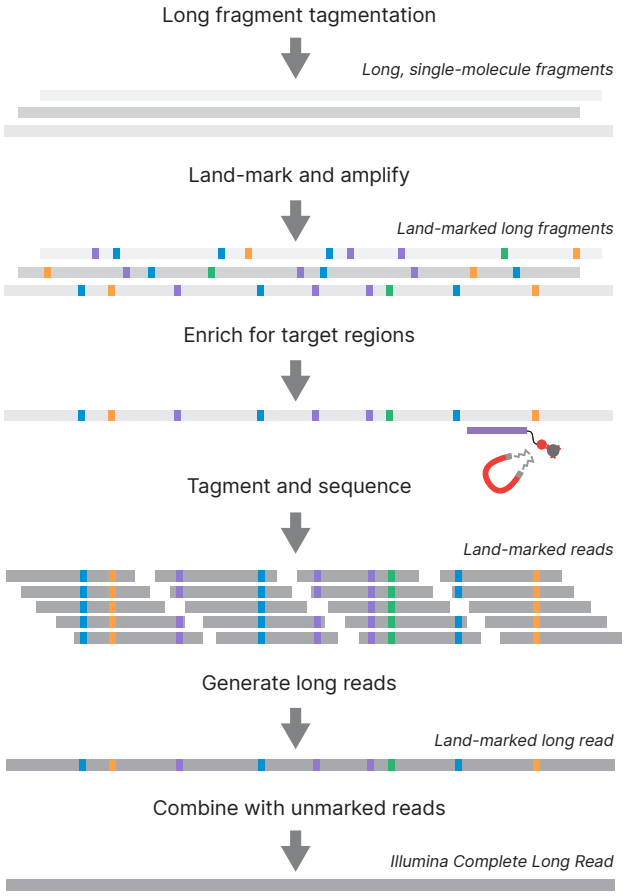


Figure 2: How the Illumina Complete Long Reads assay works—The assay uses tagmentation to make long DNA fragments (> 10 kb), eliminating the need for shearing or size selection. Long fragments are "land-marked" at the single-molecule scale to capture and preserve long-read information within the fragment (without complex barcodes or adapters). Land-marked long fragments are enriched for target regions using a hybrid–capture probe panel. Then the enriched long fragments are tagmented one more time into a sequencing-ready library. Analysis generates long reads and combines the data with a standard, unmarked WGS library (from the same sample, sequenced separately) to produce highly accurate Illumina Complete Long Reads.

Table 1: Recommended sample throughput to generate 30× final coverage for Illumina Complete Long Read Prep with Enrichment, Human[a,b,c]

	Illumina Human Comprehensive Panel	Large-sized panel	Medium-sized panel	Small-sized panel		
Target region size[d]	> 95 Mb	~20 Mb	~10 Mb	~5 Mb		
Recommended data per sample[a]	90–120 Gb	30–60 Gb	15–30 Gb	7.5–15 Gb		
300-cycle reagent kits	No. of samples per flow cell[e]				Output per flow cell	Run time
NovaSeq 6000 SP Reagents	2	8	16	32	~250 Gb	~25 hr
NovaSeq 6000 S1 Reagents	4	16	32	66	~500 Gb	~25 hr
NovaSeq 6000 S2 Reagents	10	40	82	166	~1.25 Tb	~36 hr
NovaSeq 6000 S4 Reagents	24	100	200	400[f]	~3 Tb	~44 hr
NovaSeq X Series 1.5B Reagents	4	16	32	66	~500 Gb	~21 hr
NovaSeq X Series 10B Reagents	24	100	200	400[f]	~3 Tb	~25 hr
NovaSeq X Series 25B Reagents	64	266	532[f]	1066[f]	~8 Tb	~48 hr

a. Requires 2 × 150 bp sequencing run and 5M–10M paired-end reads (~1.5–3 Gb data) per Mb target region, generating approximately 30× final coverage of Illumina Complete Long Reads. Custom panel data requirements per sample are only a recommended starting point. Users may optimize alloted data based on panel performance.
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; FASTQ files from a previously run sample can be used.
c. Sequencing Illumina Complete Long Reads 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.
d. Target region size is the sum of padded probe location lengths, merged where they overlap.
e. Example sample throughput numbers calculated for 5M paired-end reads (1.5 Gb data) per Mb target region.
f. A maximum of 384 unique dual indexes is available. For NovaSeq X Series, independent lane loading allows for multiplexing of more samples. Use the NovaSeq 6000 Xp Workflow for independent lane loading on the NovaSeq 6000 System.

These panels target challenging medically relevant genes (CMRG),[10] genes commonly targeted by pharmacogenetic (PGx) testing assays,[11–13] genes on the American College of Medical Genetics and Genomics (ACMG) secondary findings list (ACMG SF v3.1),[14] or the full major histocompatibility complex (MHC) region.[15]

Custom panels

Users can leverage their own knowledge to create a unique panel focused on a subset of genes. The DesignStudio tool supports custom panel design with an algorithm tuned for long-fragment enrichment. Third-party oligo panels are also compatible with Illumina Complete Long Read Prep with Enrichment, Human.

Scalable, high-throughput workflow

The Illumina Complete Long Read Prep with Enrichment, Human workflow is easy to automate, requires only standard lab equipment, and is highly scalable to support comprehensive WGS for more samples (Figure 1). The simple library prep protocol takes approximately 6.5 hours on day one (with ~3 hours hands-on time), followed by an overnight hybridization reaction, and 5.5 hours on day two (with ~3 hours hands-on time).

Use the NovaSeq X Series 10B or 25B flow cells for greater throughput for larger cohorts and better cost for larger panels (Table 1). With the NovaSeq X Plus System, users can generate up to 15,000 enhanced, high-accuracy genomes per year. Use lower throughput consumables like the NovaSeq X Series 1.5B flow cell for small panels and to reduce batching requirements.

Table 2: Hybrid-capture probe panels designed for Illumina Complete Long Read Prep with Enrichment, Human

Panel	Genes or regions targeted	Panel size
Illumina Human Comprehensive Panel[a]	Low-coverage spots across > 6500 protein-coding genes[9]	> 95 Mb
CMRG panel[b]	391 medically relevant genes known to be challenging to resolve with short reads[10]	22.5 Mb
PGx panel[b]	98 genes commonly targeted by pharmacogenetic testing assays[11–13]	8.1 Mb
ACMG panel[b]	78 unique genes from the ACMG secondary findings list (ACMG SF v3.1)[14]	7 Mb
MHC panel[b]	The full MHC region (> 140 genes) in the GRCh38.p14 assembly[15]	4.9 Mb

a. Premanufactured, optimized panel.
b. Predesigned and custom panels available in DesignStudio software. CMRG, challenging medically relevant genes; PGx, pharmacogenomics; ACMG, American College of Medical Genetics and Genomics; MHC, major histocompatibility complex.

Streamlined, comprehensive analysis

Data analysis for Illumina Complete Long Read Prep with Enrichment, Human is available as a BaseSpace™ Sequence Hub app or through Illumina Connected Analytics. The single DRAGEN pipeline analyzes both short and long reads to provide comprehensive WGS results. Results are merged into a single set of output files including DRAGEN targeted callers.[16]

Access highly accurate WGS with enriched long-read data

Illumina Complete Long Read Prep with Enrichment, Human data demonstrate coverage and accuracy improvements in targeted challenging regions compared to standard short-read WGS alone (Figure 3, Figure 4). Targeted Illumina Complete Long Reads also successfully resolve haplotypes of full genes and can phase large blocks of highly polymorphic regions like the MHC (Figure 5).

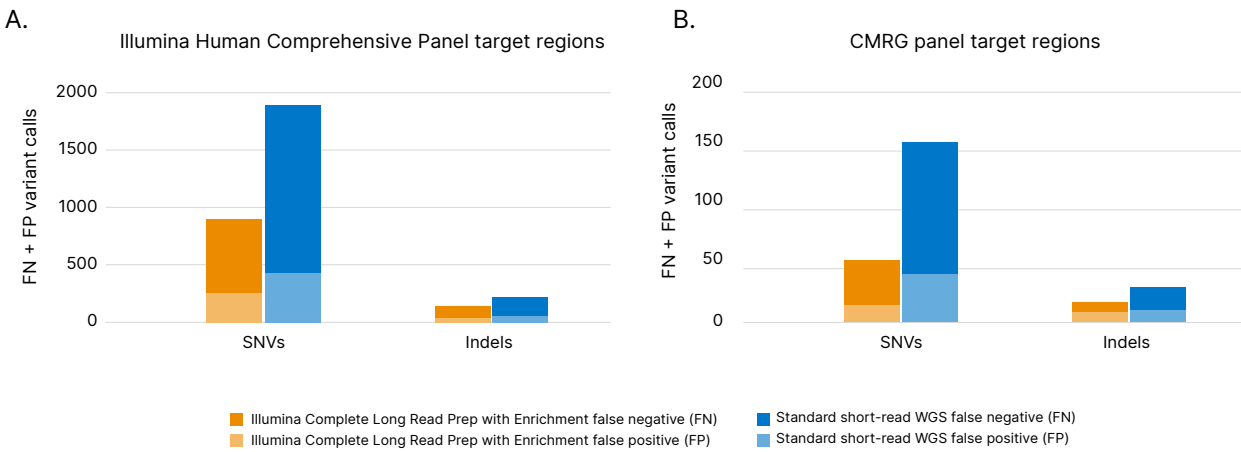


Figure 3: Targeted long reads to enhance variant calling accuracy in challenging regions—False negative (FN) plus false positive (FP) variant calls for SNVs and indels in HG002 genic regions targeted by the (A) Illumina Human Comprehensive Panel or (B) CMRG panel, using Illumina Complete Long Read Prep with Enrichment (orange) compared to standard short-read WGS (blue).

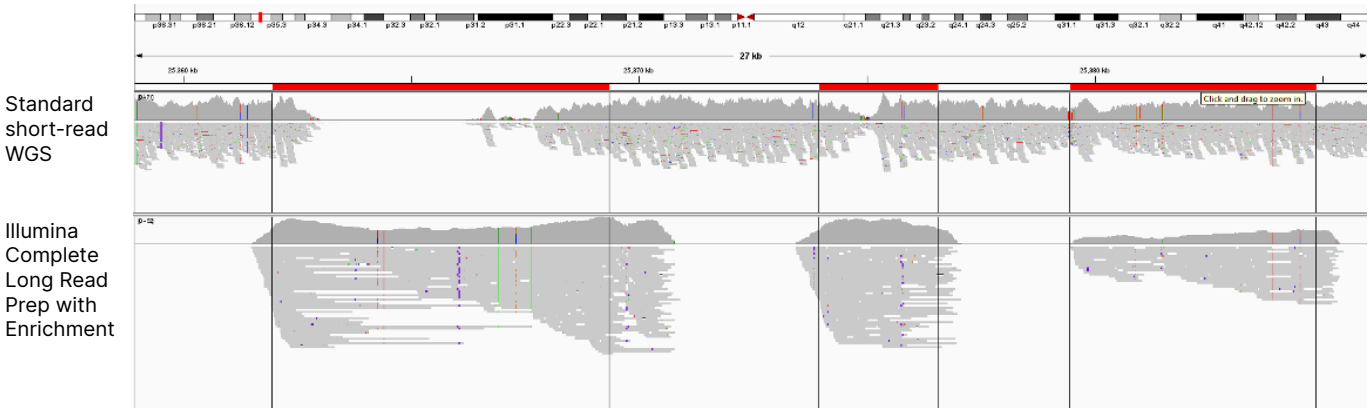
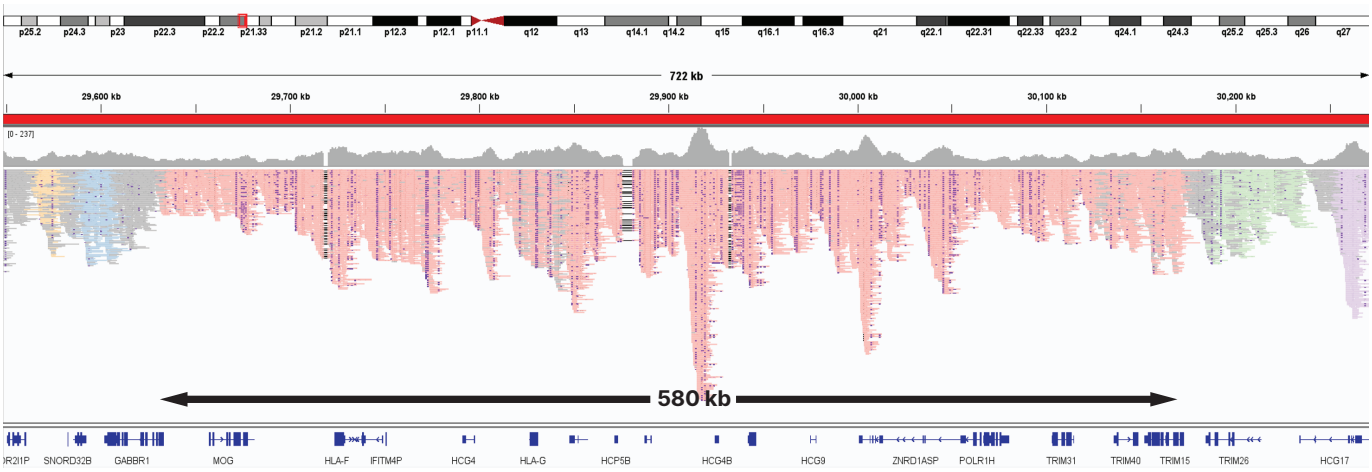


Figure 4: Achieve cost-effective human whole-genome coverage of challenging regions—Illumina Complete Long Read Prep with Enrichment, Human helps enhance coverage in challenging genic regions like those in the *RHCE* gene to complement standard short-read human WGS. Integrative Genomics Viewer (IGV) plot of *RHCE* sequenced using standard short-read WGS (top) and Illumina Complete Long Read Prep with Enrichment (bottom). Target regions noted in red.

A.



B.

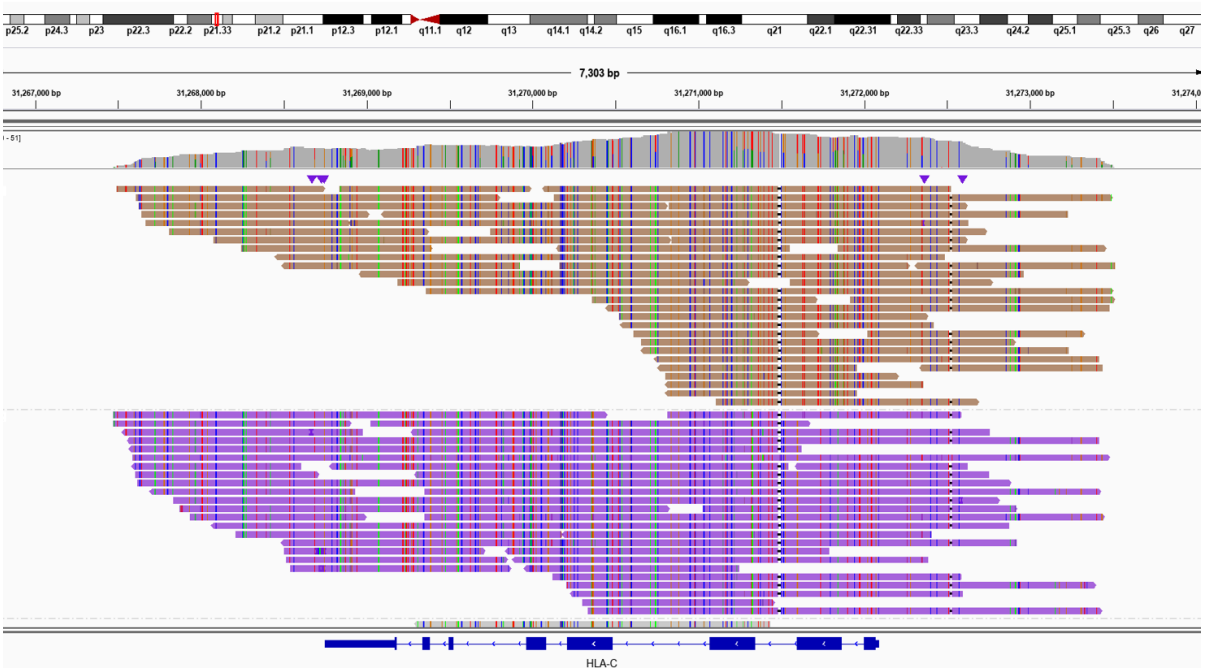


Figure 5: Targeted long reads help resolve haplotypes in polymorphic genes—IGV plots from long-read sequencing using Illumina Complete Long Read Prep with Enrichment, Human. (A) Phasing over a 722 kb region in the MHC locus. A 580 kb region (pink) is encapsulated in one phase block. (B) *HLA-C* gene is fully phased. Reads are separated by haplotype.

Summary

Illumina Complete Long Read Prep with Enrichment, Human is a flexible, cost-effective solution that complements proven Illumina WGS and focuses on long reads where they provide greatest value. Illumina Complete Long Reads makes comprehensive WGS easily accessible for genomics labs by enabling both long and short reads on the same instrument with a single DRAGEN analysis pipeline. The streamlined, familiar workflow delivers the most scalable and accurate whole genome assay on the market.

Learn more

Illumina Complete Long Read Prep with Enrichment, Human

Human Comprehensive Panel

Long-read sequencing technology

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. Published 2020. Accessed January 12, 2023.
2. 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.
3. Pacific Biosciences. Preparing DNA for PacBio HiFi sequencing—Extraction and quality control. pacb.com/literature/technical-note-preparing-dna-for-pacbio-hifi-sequencing-extraction-and-quality-control/. Published 2022. Accessed January 12, 2023.
4. 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 October 5, 2023.
5. Oxford Nanopore Technologies. Ligation Sequencing Kit. https://store.nanoporetech.com/us/productDetail/?id=ligation-sequencing-kit-v14. Accessed October 5, 2023.
6. Pacific Biosciences. Low Yield Troubleshooting Guide. pacb.com/wp-content/uploads/Guide-Low-Yield-Troubleshooting.pdf. Published 2018. Accessed January 12, 2023.
7. Illumina. Illumina Complete Long Read Prep, Human data sheet. illumina.com/content/dam/illumina/gcs/assembled-assets/marketing-literature/illumina-long-read-prep-human-data-sheet-m-gl-01420/illumina-long-read-prep-data-sheet-m-gl-01420.pdf. Published 2022. Accessed September 22, 2023.
8. Illumina. Comprehensive whole-genome sequencing with Illumina Complete Long Read Prep, Human technical note. illumina.com/content/dam/illumina/gcs/assembled-assets/marketing-literature/illumina-long-read-prep-human-tech-note-m-gl-01421/ilmn-long-read-hu-tech-note-m-gl-01421.pdf. Published 2022. Accessed September 22, 2023
9. Bekritsky MA, Colombo C, Eberle MA. Identifying genomic regions with high quality single nucleotide variant calling. illumina.com/science/genomics-research/articles/identifying-genomic-regions-with-high-quality-single-nucleotide-.html. Published 2021. Accessed August 30, 2023.
10. Wagner J, Olson ND, Harris L, et al. Curated variation benchmarks for challenging medically relevant autosomal genes. *Nat Biotechnol*. 2022;40(5):672-680. doi:10.1038/s41587-021-01158-1
11. PharmGKB. VIPs: Very Important Pharmacogenes. pharmgkb.org/vips. Accessed September 22, 2023.
12. National Library of Medicine. GTR: Genetic Testing Registry. Precision HealthPGx Panel (25 Genes). ncbi.nlm.nih.gov/gtr/tests/593428/. Updated November 29, 2022. Accessed September 22, 2023.
13. Pratt VM, Everts RE, Aggarwal P, et al. Characterization of 137 Genomic DNA Reference Materials for 28 Pharmacogenetic Genes: A GeT-RM Collaborative Project. *J Mol Diagn*. 2016;18(1):109-123. doi:10.1016/j.jmoldx.2015.08.005
14. Miller DT, Lee K, Abul-Husn NS, et al. ACMG SF v3.1 list for reporting of secondary findings in clinical exome and genome sequencing: A policy statement of the American College of Medical Genetics and Genomics (ACMG). *Genet Med*. 2022;24(7):1407-1414. doi:10.1016/j.gim.2022.04.006
15. Kulski JK, Suzuki S, Shiina T. Human leukocyte antigen super-locus: nexus of genomic supergenes, SNPs, indels, transcripts, and haplotypes. *Hum Genome Var*. 2022;9(1):49. doi:10.1038/s41439-022-00226-5
16. Roessler K. Illumina Complete Long Reads software analysis workflow for human WGS. illumina.com/science/genomics-research/articles/complete-long-read-software-analysis.html. Published 2023. Accessed September 22, 2023.

Ordering information

Product	Catalog no.
Illumina Complete Long Read Prep with Enrichment, Human (24 samples)	20113832
Illumina Complete Long Read Prep with Enrichment, Human (96 samples)	20113833
Illumina Complete Long Read Prep with Enrichment, Human Comprehensive Panel (24 samples)	20113834
Illumina Complete Long Read Prep with Enrichment, Human Comprehensive Panel (96 samples)	20113835
Illumina Human Comprehensive Panel (24 samples)	20113836
Illumina Human Comprehensive Panel (96 samples)	20113837
Illumina Custom Enrichment Panel v2 (32 µl, 120 bp)	20073953
Illumina Custom Enrichment Panel v2 (384 µl, 120 bp)	20073952
Illumina Custom Enrichment Panel v2 (1536 µl, 120 bp)	20111339

Ordering information

Product	Catalog no.
Illumina Unique Dual Indexes, LT (48 indexes, 48 samples)	20098166
Illumina DNA/RNA UD Indexes Set A, Tagmentation (96 indexes, 96 samples)	20091654
Illumina DNA/RNA UD Indexes Set B, Tagmentation (96 indexes, 96 samples)	20091656
Illumina DNA/RNA UD Indexes Set C, Tagmentation (96 indexes, 96 samples)	20091658
Illumina DNA/RNA UD Indexes Set D, Tagmentation (96 indexes, 96 samples)	20091660
Illumina Analytics - 1 iCredit	20042038

illumina®

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

M-GL-02188 v1.0