DATA SHEET

# Illumina Human Comprehensive Panel

## Enhanced coverage of challenging regions in protein-coding genes with Illumina Complete Long Read Prep with Enrichment, Human

illumina®

# Introduction

When performing human whole-genome sequencing (WGS), there may be a small fraction of genic regions that are difficult to map with short reads alone. Challenging regions include highly homologous or repetitive regions, complex structural variants, pseudogenes, and large insertion–deletions (indels). In these cases, long-read sequencing can complement standard short-read WGS data to help deliver enhanced coverage to resolve challenging regions and make it easier to phase variants and call haplotypes.

Illumina Complete Long Reads technology uses a standard next-generation sequencing (NGS) workflow to generate contiguous long-read sequences on Illumina sequencing systems (Figure 1). Illumina Complete Long Read Prep with Enrichment and the Human Comprehensive Panel offer a cost-effective, scalable approach to help researchers resolve known challenging regions of the genome. Targeted long reads are analyzed with standard high-accuracy WGS data to enable mapping of more regions.

## Comprehensive and optimized probe panel

The Illumina Human Comprehensive Panel is an optimized hybrid–capture probe panel designed to address genic regions that can benefit from increased mappability of longer reads. The panel targets low-mappability regions across > 6500 protein-coding genes (Table 1)[1] to enable complementary long reads that enhance the resolution of areas that are challenging to map using standard short-read WGS methods.

Table 1: Human Comprehensive Panel parameters[a]

| | |
|---|---|
| System | NovaSeq X Series<br>NovaSeq 6000 System |
| Target region size | > 95 Mb |
| Sequencing output per sample[b] | 90–120 Gb |
| No. of probes | ~40K |
| No. of genes targeted | > 6500 |
| Sample type | Genomic DNA |
| Recommended DNA input | 50 ng |
| Total library prep time | ~2 days |
| Hands-on time | ~6 hr |
| Multiplexing | Up to 64 samples |
| N50 | 6.3 kb |
| Phase block N50 | 15.6 kb |
| Uniformity | 95% |
| Padded read enrichment (PRE)[c] | 83% |
| % Heterozygous SNV phased[d] | 98% |

a. Data generated using 50 ng HG002 genomic DNA (Corielle, Catalog no. NA24385). Performance may vary with DNA input and sample quality.
b. Requires 2 × 150 bp sequencing run, generating approximately 30× final coverage of Illumina Complete Long Reads.
c. Uniformity of coverage calculated as % > 0.2 * mean.
d. PRE calculated as 100 * (padded target aligned reads / total aligned reads).
e. SNV, single nucleotide variants.

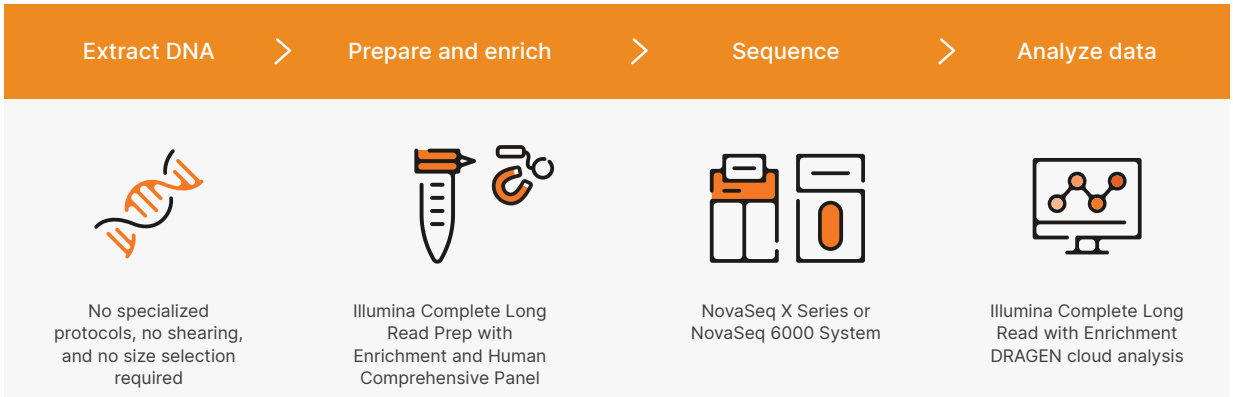


Figure 1: Part of an integrated workflow—Access cost-effective, targeted long-read WGS data using a scalable, optimized library prep with enrichment protocol, proven Illumina sequencing chemistry, and DRAGEN secondary analysis. Requires ≥ 30× standard short-read WGS data from the same sample for analysis.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

In developing this panel, Illumina considered the full set of over 20,000 protein-coding genes, including introns and untranslated regions (UTRs), and targeted 70% of all difficult-to-map base pairs across these regions. Genes that are comprehensively covered by short reads alone were excluded. This panel has been wet-lab tested and optimized to maximize design and sequencing efficiency and performance.

## Highly scalable sequencing and analysis

The Illumina Complete Long Read Prep with Enrichment workflow is highly scalable and easy to automate to support comprehensive WGS for a larger number of samples (Table 2). On the NovaSeq™ X Plus System, users can generate up to 15,000 high-accuracy genomes per year.* To reduce batching requirements, sequence smaller sample numbers on lower throughput consumables like the NovaSeq X 1.5B flow cell. Illumina Complete Long Read Prep with Enrichment and Human Comprehensive Panel can be used to augment existing WGS data sets as a reflex tool for broader variant detection.

Data analysis for Illumina Complete Long Read Prep with Enrichment, Human is available as a BaseSpace™ Sequence Hub app or through Illumina Connected Analytics. Data is streamed to the cloud and analyzed together with standard short-read ≥ 30× WGS data from the same sample (generated previously or in parallel). The DRAGEN™ pipeline merges results into a single set of output files including DRAGEN targeted callers.[3]

## Cost-effective, enhanced coverage

Illumina Complete Long Read Prep with Enrichment and the Human Comprehensive Panel improve variant calling accuracy and enhance coverage across target regions, compared to standard short-read WGS (Figure 2, Figure 3, Table 3). Targeted long reads using the Human Comprehensive Panel deliver comprehensive whole-genome accuracy with an F1 score (SNVs + indels) of 99.87%,[2] a substantial improvement over short reads alone. This provides performance comparable with Illumina Complete Long Read Prep, Human and other long-read whole genomes in protein-coding genes, yet with lower cost and higher throughput (Table 3).

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

Table 2: Recommended sample throughput to generate 30× final coverage for Human Comprehensive Panel and Illumina Complete Long Read Prep with Enrichment, Human[a,b,c,d]

| | NovaSeq 6000 | | | | NovaSeq X | | |
|---|---|---|---|---|---|---|---|
| 300-cycle reagent kits | SP | S1 | S2 | S4 | 1.5B | 10B | 25B |
| Samples per flow cell | 2 | 4 | 10 | 24 | 4 | 24 | 64 |
| Output per flow cell | ~250 Gb | ~500 Gb | ~1.25 Tb | ~3 Tb | ~500 Gb | ~3 Tb | ~8 Tb |
| Run time | ~25 hr | ~25 hr | ~36 h | ~44 hr | ~21 hr | ~25 hr | ~48 h |

a. Human Comprehensive Panel target region size is > 95 Mb and requires 90–120 Gb sequencing output per sample.
b. Requires 2 × 150 bp sequencing run, generating approximately 30× final coverage of Illumina Complete Long Reads.
c. 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.
d. 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.

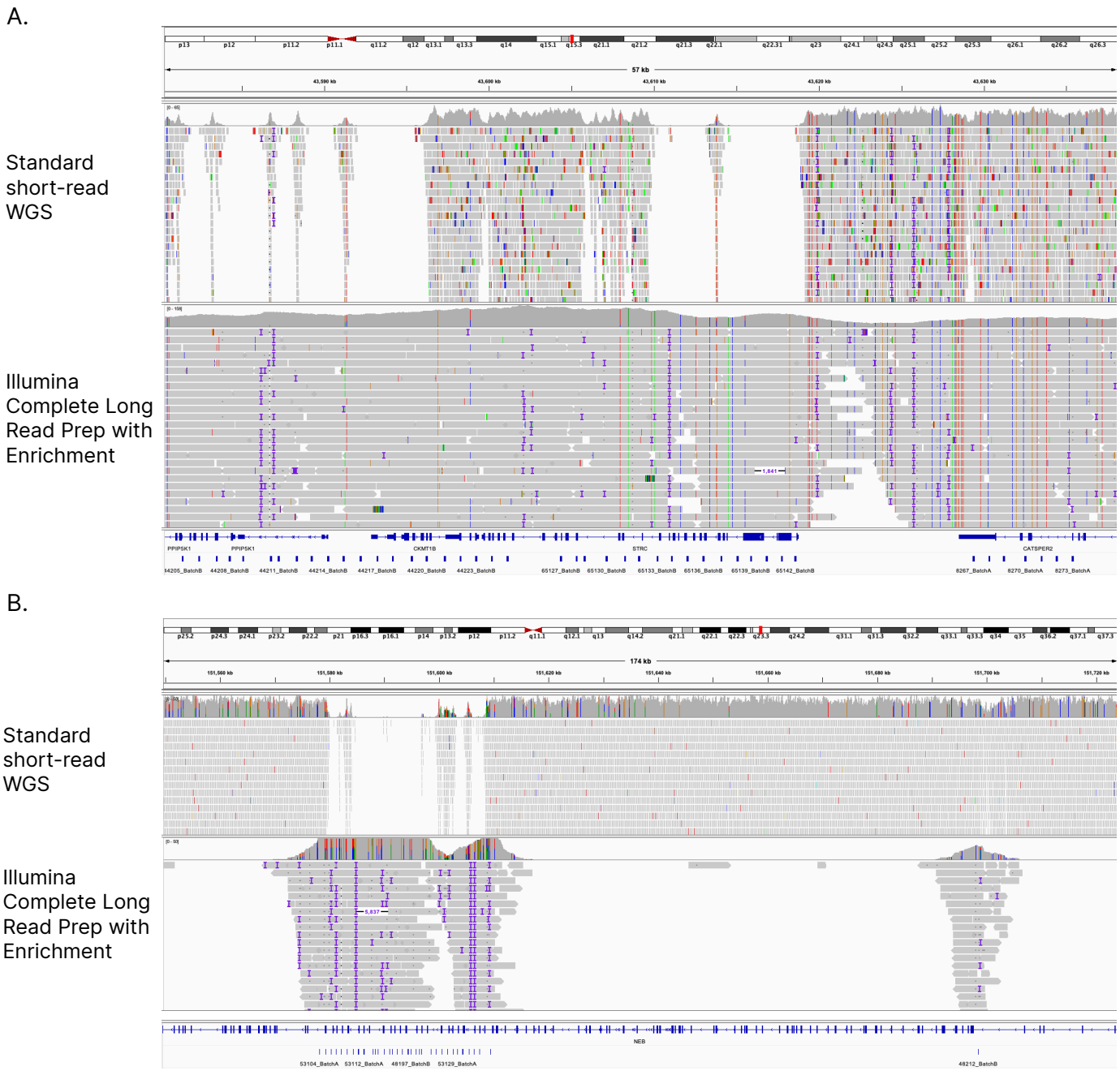


Figure 2: Resolve difficult-to-map regions using targeted long reads with the Human Comprehensive Panel—Illumina Complete Long Read Prep with Enrichment, Human and the Human Comprehensive Panel help enhance coverage in challenging genic regions to complement standard short-read human WGS. Integrative Genomics Viewer (IGV) plots of (A) *STRC* and (B) *NEB* using standard short-read WGS and Illumina Complete Long Reads with Enrichment.

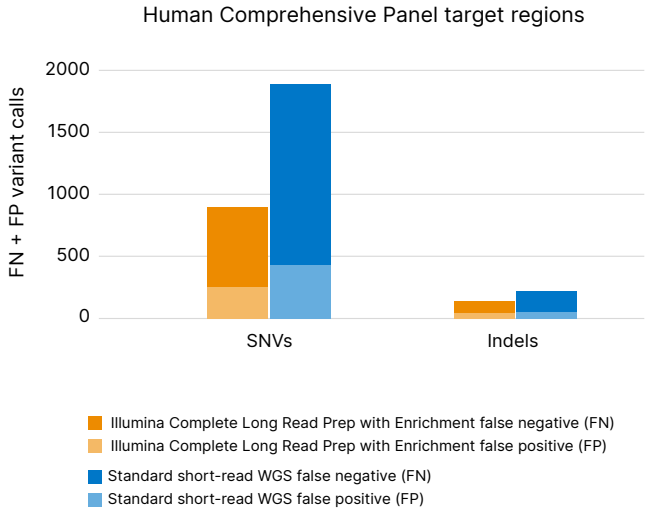


Figure 3: Targeted long reads to improve variant calling accuracy—False negative (FN) plus false positive (FP) SNV and indel variant calls in genic regions targeted by the Human Comprehensive Panel, using Illumina Complete Long Read Prep with Enrichment and Human Comprehensive Panel (orange) compared to standard short-read WGS (blue).

Table 2: Human Comprehensive Panel performance

| Accuracy for Human Comprehensive Panel target regions[a] | | |
|---|---|---|
| | Illumina Complete Long Read Prep with Enrichment and Human Comprehensive Panel[b] | Standard short-read WGS[c] |
| F1 score (SNVs) | 99.54% | 99.04% |
| F1 score (Indels) | 99.47% | 99.15% |
| F1 score (SVs) | 80.43% | 60.18% |

| Accuracy for whole genome, all benchmark regions[d] F1 score (SNVs + indels) | | |
|---|---|---|
| Illumina Complete Long Read Prep with Enrichment and Human Comprehensive Panel[b] | Illumina Complete Long Read Prep, Human[e] | Standard short-read WGS[c] |
| 99.87% | 99.90% | 99.84% |

a. Autosomes only. SNV, single nucleotide variants; SV, structural variants.
b. Illumina Complete Long Read Prep with Enrichment, Human Comprehensive Panel + Illumina DNA PCR-Free Prep + NovaSeq 6000 System + DRAGEN v4.2.
c. Illumina DNA PCR-Free Prep + NovaSeq 6000 System + DRAGEN v4.2.
d. As measured against all participating solutions using PrecisionFDA Truth Challenge v2 Benchmark Data,[2] internal data on file for DRAGEN v4.
f. Illumina Complete Long Read Prep, Human + Illumina DNA PCR-Free Prep + NovaSeq 6000 System + DRAGEN v4.2.

## Summary

The Illumina Human Comprehensive Panel enables high-accuracy long-read coverage of the small portion of genic regions that are challenging to map with short reads alone. IIllumina Complete Long Read Prep with Enrichment and Human Comprehensive Panel offer an optimized, cost-effective whole-genome assay that complements Illumina WGS and focuses on long reads where they provide greatest value with a full workflow solution.

## Learn more

Illumina Complete Long Read Prep with Enrichment, Human

Human Comprehensive Panel

Long-read sequencing technology

## References

1. Bekritsky MA, 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.
2. PrecisionFDA. Truth Challenge V2: Calling Variants from Short and Long Reads in Difficult-to-Map Regions. precision.fda.gov/challenges/10. Accessed October 2, 2023.
3. 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 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-02191 v1.0