illumina®

TruSight™ Exome Library Preparation Kit

A focused subset of the exome targeting genes with disease-causing mutations delivered on proven next-generation sequencing technology.

Highlights

- **Focused Content**
  Content focused on disease-causing mutations as curated by the Human Genome Mutation Database (HGMD)
- **Low Input DNA Requirement**
  Excellent data quality with as little as 50 ng DNA to preserve precious samples
- **Fast, Simple Workflow**
  Library preparation and enrichment completed in 1.5 days

Introduction

Consisting of coding regions of expressed genes, the exome represents the most functionally relevant part of the genome. Even so, it is only 1–2% of the entire genome. Focusing sequencing efforts on this small portion has proven to be a time- and cost-efficient manner for identifying variants linked with diseases and other health conditions. TruSight Exome content sets simplify and further speed up this discovery and recognition process by providing pre-designed, ready-to-use oligos specifically targeting genes with known associations to inherited diseases. The content set is compatible with TruSight Rapid Capture that takes advantage of Nextera® Rapid Capture technology to offer a single, integrated library preparation and enrichment workflow that can be completed in just 1.5 days (Figure 1). Delivering excellent data quality from low sample input (50 ng), TruSight Exome and TruSight Rapid Capture kits enable researchers to access precious samples, while retaining sufficient material for future analyses.

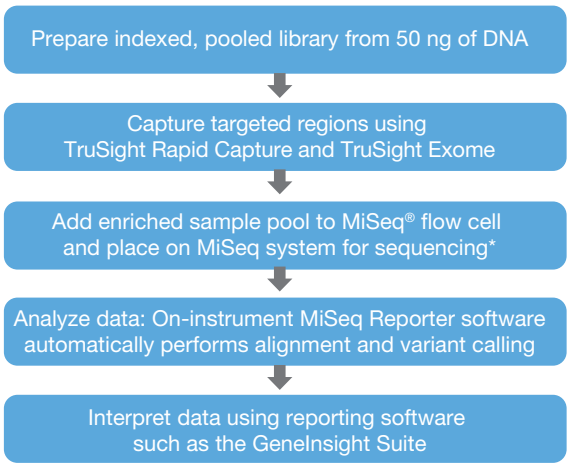
Content Design Strategy

Developed by Illumina, the TruSight Exome content set contains a subset of the known genes in the full exome. It is focused on disease-causing mutations as curated by the Human Genome Mutation Database (HGMD) and shown to be important in specific inherited conditions[1].

Superior Coverage

The TruSight Exome content set features a highly optimized probe set that supports discovery of a large number of variants. Starting with only 50 ng of DNA input, the content set delivers comprehensive coverage of the targeted exomic sequences. The content set includes > 50,000 80-mer probes, each constructed against the human NCBI37/hg19 reference genome. The probe set was designed to enrich for ~40,000 exons, spanning 2,761 genes of interest (Table 1).

Figure 1: Simple, Integrated Workflow



TruSight Exome is compatible with the rapid capture enrichment method, which integrates library preparation and enrichment steps to offer a fast, streamlined, and optimized workflow, delivering fully enriched libraries for up to 96 samples in just 1.5 days.

* For higher throughput, use HiSeq flow cells and the HiSeq system.

TruSight Exome targets 7.00 Mb of the human genome. The 80-mer probes target libraries of approximately 500 bp (insert size of 300 bp), enriching 350–650 bases centered symmetrically around the midpoint of the probe (Figure 2)[2]. This means that, in addition to comprehensive coverage of the major exon regions, the kit provides broad coverage of non-coding DNA in exon-flanking regions (splice sites). By focusing on this region, or subset, of the exome with known associations to inherited disease (as indicated by the HGMD), TruSight Exome enables labs to detect variants that affect gene function more efficiently than whole-genome or whole-exome sequencing[3].

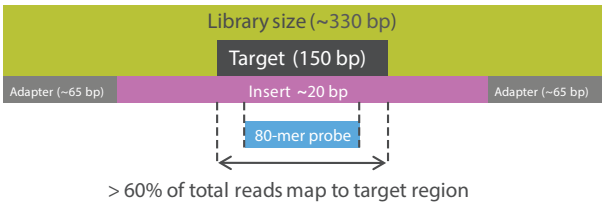
Integrated Library Preparation and Enrichment Workflow

TruSight Exome and TruSight Rapid Capture leverage the speed of Nextera library preparation technology. By eliminating the need for mechanical DNA fragmentation and introducing a unique multiplex pre-enrichment sample pooling, the TruSight Rapid Capture method reduces hands-on time for a high-throughput workflow that saves at least one full day over all other currently available enrichment workflows. Furthermore, master-mixed reagents are coupled with a plate-based

Table 1: Coverage Details

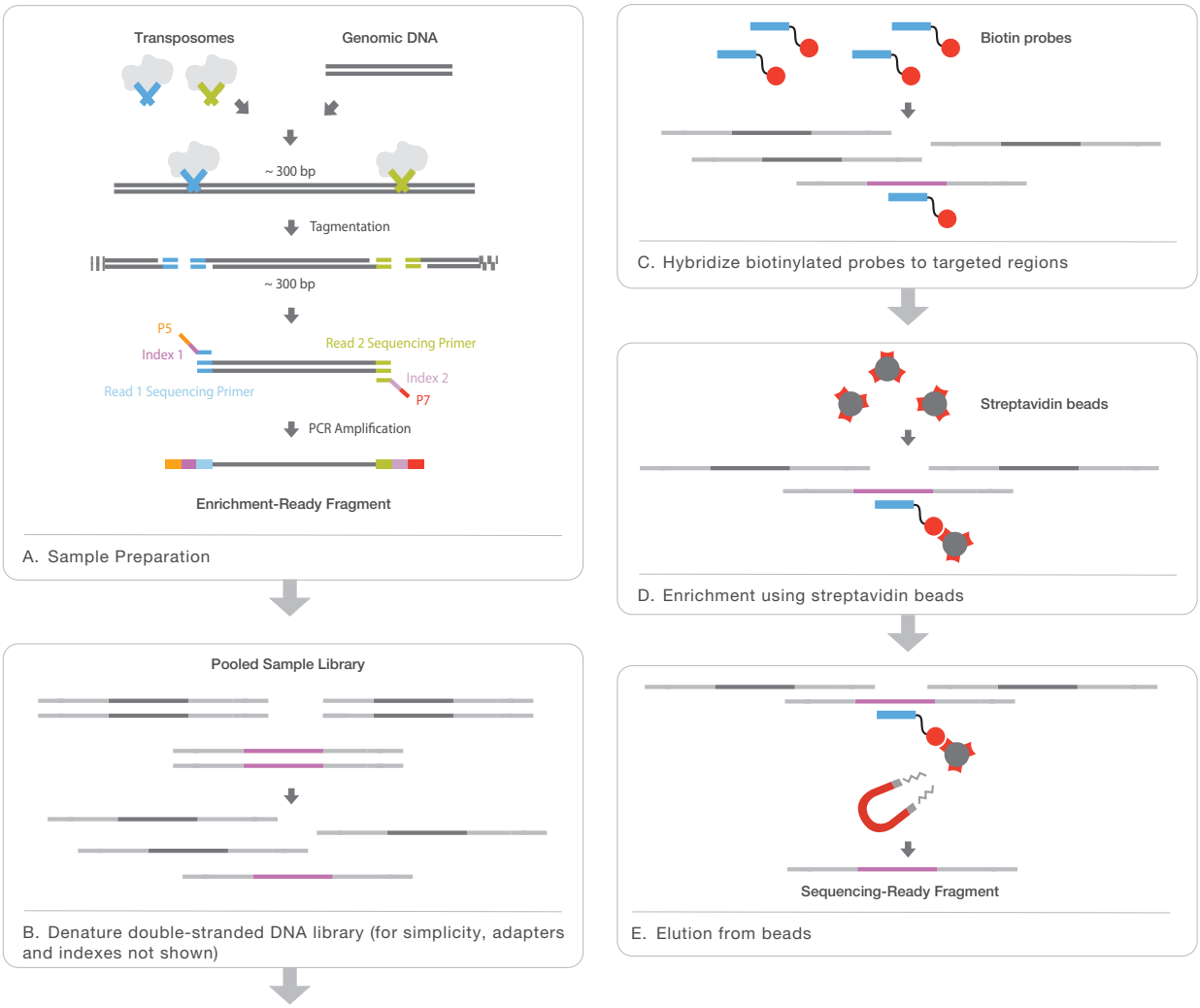
Cumulative target region size	7.00 Mb
Number of target genes	2,761
Number of target exons	~40,000
Probe size	80-mer
Number of probes	> 50,000
Recommended mean coverage	100×
Target minimum coverage	12×
Percent exons covered based on coverage metrics	≥ 95%

Figure 2: Probe Footprint



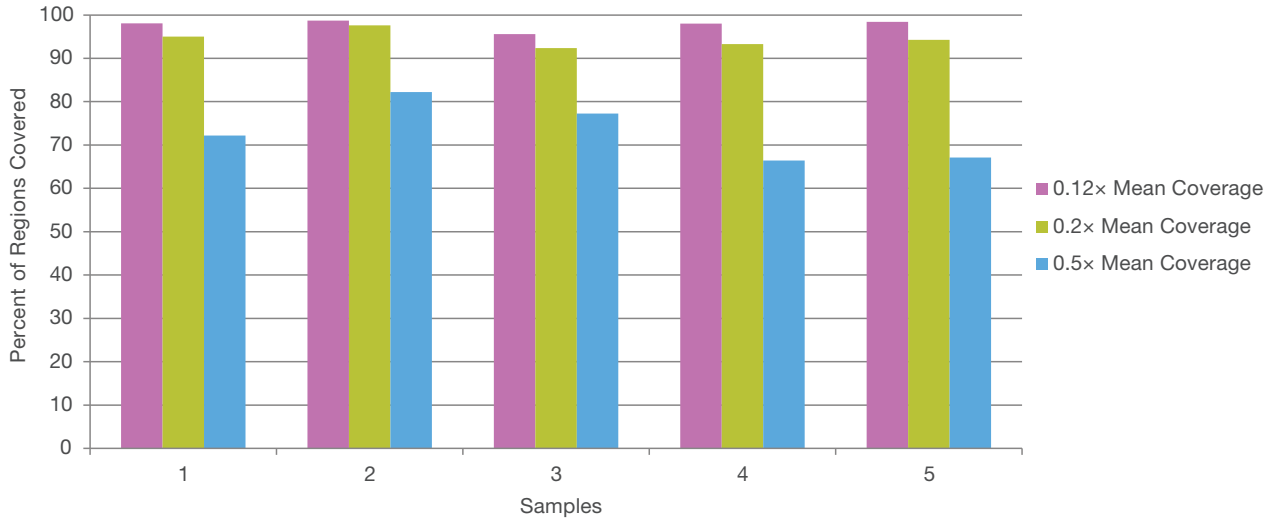
With an approximately 500 bp DNA library (insert size of 300 bp), the probe will enrich 350–650 bp centered around its midpoint.

Figure 3: Integrated TruSight Rapid Capture Workflow



The TruSight Rapid Capture workflow provides a fast, simple method for isolating the genes targeted using TruSight Exome. The streamlined, automation-friendly workflow combines library preparation and exome enrichment steps, and can be easily completed in 1.5 days with minimum hands-on time.

Figure 4: High Coverage Uniformity



Coverage uniformity is given for five samples with respect to the percentage of targeted regions at varying mean normalized read depths. The five samples were prepared and then enriched using the TruSight Rapid Capture Kit along with the TruSight Exome content set. Samples were individually sequenced across a MiSeq standard flow cell, generating mean read depths of 98–175× (varying for each sample). Over 95% of bases (7 Mb) were covered at 0.12× mean coverage for each sample.

protocol for simultaneous processing of up to 24 enrichment reactions (288 total samples).

Flexible kit configurations enable labs to readily meet their sample throughput needs. For those requiring higher throughput, kit reagent volumes are optimized for liquid handlers to make an automation-friendly workflow. TruSight Rapid Capture kits supporting lower throughput options are also available, allowing labs to cost-effectively run samples immediately instead of waiting to batch.

Following the TruSight workflow, the process starts with rapid Nextera-based library prep to convert input genomic DNA into adapter-tagged libraries (Figure 3A). This rapid prep requires only 50 ng of input DNA and takes less than 3 hours for a plate of 96 samples. Nextera tagmentation of DNA simultaneously fragments and tags DNA without the need for mechanical shearing. Integrated sample barcodes then allow the pooling of up to 96 samples for a single Rapid Capture pull down. Next, libraries are denatured into single-stranded DNA (Figure 3B) and biotin-labeled probes specific to the targeted region are used for the Rapid Capture hybridization (Figure 3C). The pool is enriched for the desired regions by adding streptavidin beads that bind to the biotinylated probes (Figure 3D). Biotinylated DNA fragments bound to the streptavidin beads are magnetically pulled down from the solution (Figure 3E). The enriched DNA fragments are then eluted from the beads and hybridized for a second Rapid Capture. This entire process is completed in only 1.5 days, enabling a single researcher to efficiently process up to 288 samples at one time—all without automation.

Data Analysis

Sequence data generated from TruSight Exome–enriched libraries are analyzed by the on-instrument MiSeq Reporter (MSR) software. After demultiplexing and FASTQ file generation, the software uses the Burrows-Wheeler Aligner (BWA) to align the reads against the hg19 homo sapiens reference genome to create BAM files. The Genome Analysis Toolkit (GATK) is then used to perform variant analysis for the target regions specified in the manifest file. The output of GATK are VCF files, which are text files that contain SNPs.

High Data Quality

With TruSight Exome and TruSight Rapid Capture, researchers can be confident in the quality of sequencing data generated from pooled multisample libraries. Each sample is sequenced with high coverage uniformity across the target region, with 95% of exons covered at a minimum coverage of 12× (Figure 4). This uniformity applies to smaller exons (< 150 bp) as well as long coding exons.

Flexible Exome Options

With studies showing that many rare disease–causing variants lie within the 1–2% of the coding, or exonic, region of the genome, targeted exome sequencing is becoming more widely used. Illumina offers a number of options for exome enrichment studies, enabling labs to obtain the coverage and scalability they need in the most time- and cost-efficient manner (Table 2). Learn more about Illumina exome enrichment options at www.illumina.com.

Summary

TruSight Exome enables labs to access the most promising regions of the genome for targeting rare diseases. The optimized probe set provides comprehensive coverage of the targeted regions with high coverage uniformity for identifying a large number of variants. Combining this content with the TruSight Rapid Capture method enables a fast, easy workflow, requiring low sample DNA input and providing a highly efficient resequencing solution to accelerate detection of variants associated with inherited disease.

Table 2: Flexible Exome Options

	TruSight Exome	Nextera Rapid Capture Exome	Nextera Rapid Capture Extended Exome
Enrichment method	Nextera rapid capture	Nextera rapid capture	Nextera rapid capture
Target region size	7 Mb	45 Mb	62 Mb
Description of targets	Covers coding exons of genes underlying or associated with human inherited disease as described in the Human Gene Mutation Database (HGMD)	Covers > 98% of RefSeq, Ensembl CCDS, and GENCODE v12 coding exons	Covers > 97% of Ensembl CCDS, > 96% of RefSeq, and > 93% of Encode/Gencode coding exons; also includes > 88% of RefSeq 5' and 3' UTRs and > 77% of predicted microRNA targets
Uniformity (% of exons within 0.2× of mean coverage)	> 95%	> 80%	> 80%
Read length recommendation	2 × 150	2 × 75	2 × 75
DNA input	50 ng	50 ng	50 ng

Ordering Information

Product	Catalog No.	TG Catalog No.*
TruSight Exome Content Set (4 enrichments)	N/A	TG-141-1001
Rapid Capture Kits		
TruSight Rapid Capture Kit (1 index, 8 samples)	FC-140-1101	TG-140-1101
TruSight Rapid Capture Kit (2 indices, 8 samples)	FC-140-1102	TG-140-1102
TruSight Rapid Capture Kit (4 indices, 16 samples)	FC-140-1103	TG-140-1103
TruSight Rapid Capture Kit (24 indices, 48 samples)	FC-140-1104	TG-140-1104
TruSight Rapid Capture Kit (24 indices, 96 samples)	FC-140-1105	TG-140-1105
TruSight Rapid Capture Kit (96 indices, 288 samples)	FC-140-1106	TG-140-1106

* TG-labeled consumables include features intended to help customers reduce the frequency of revalidation. They are available only under supply agreement and require customers to provide a binding forecast. TruSight content sets are available for evaluation purposes prior to executing a supply agreement. Please contact your account manager for more information.

Learn More

To learn more about TruSight Exome, TruSight Rapid Capture kits, and Illumina next-generation sequencing technology, visit www.illumina.com/trusightexome.

References

1. http://www.hgmd.cf.ac.uk/ac/index.php
2. Optimizing Coverage for Targeted Resequencing Technical Note. www.illumina.com/documents/products/technotes/technote_optimizing_coverage_for_targeted_resequencing.pdf
3. Bainbridge MN, Wang M, Wu YQ, Newsham I, Muzny DM, et al. (2011) Targeted enrichment beyond the consensus coding DNA sequence exome reveals higher variant densities. Genome Biol. 12: R68.

Note regarding biomarker patents and other patents unique to specific uses of products.

Some genomic variants, including some nucleic acid sequences, and their use in specific applications may be protected by patents. Customers are advised to determine whether they are required to obtain licenses from the party that owns or controls such patents in order to use the product in customer's specific application.

Illumina • 1.800.809.4566 toll-free (U.S.) • +1.858.202.4566 tel • techsupport@illumina.com • www.illumina.com

FOR RESEARCH USE ONLY

Pub. No. 670-2012-010 Current as of 9 November 2015

illumina®