illumina®

TruSeq™ DNA Exome

A cost-effective library preparation and exome enrichment solution with exceptional accuracy.

Highlights

- **Proven TruSeq Data Quality**
  Mechanical shearing and TruSeq enrichment chemistry yield uniform coverage and ≥ 80% on-target sequencing reads

- **Cost-effective Exome Sequencing**
  Preenrichment library pooling and optimal coverage result in low-cost exome sequencing

- **Accurate, Reliable Findings**
  Push-button data analysis confidently identifies exonic variants

- **Integrated Workflow Solution**
  Comprehensive workflow streamlines exome sequencing from library prep through data analysis

Introduction

Exome sequencing has gained recognition in the scientific community as a powerful method for discovering potential causative variants for genetically driven diseases.[1–3] TruSeq DNA Exome delivers a low-cost exome sequencing solution, enabling researchers to sequence more exomes per study and accelerate their research. The kit combines proven TruSeq technology with exceptional accuracy, even for challenging samples. As part of an integrated workflow that includes library preparation, exome enrichment, sequencing, and data analysis, TruSeq DNA Exome delivers accurate variant calls and enables a deeper understanding of coding mutations.

Proven TruSeq Data Quality

Obtaining high-confidence variant calls is as much a function of sequencing accuracy as it is of high-quality library preparation and enrichment. TruSeq DNA Exome is compatible with the MiSeq™, NextSeq™, HiSeq™ and NovaSeq™ Series of sequencing systems (Table 1). These systems use sequencing by synthesis (SBS) chemistry, used to generate more than 90% of the world's sequencing data.* Illumina SBS chemistry delivers a high percentage of sequenced bases over Q30. By combining TruSeq DNA Exome with SBS chemistry, researchers can identify a high number of true coding variants and minimize false-positive and false-negative calls.

Focused Exonic Content

TruSeq DNA Exome is optimized to provide uniform and specific coverage of 45 Mb of exonic content. The probe set is designed to enrich 214,405 exons (Table 2). This focused design, paired with uniform and specific enrichment, enables comprehensive exome

*Data calculations on file. Illumina, Inc., 2015.

sequencing and reliable identification of true, coding variants.

Table 1: Throughput Comparison with TruSeq DNA Exome[a]

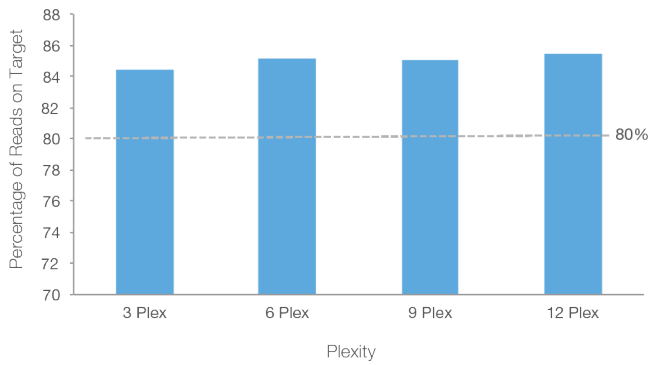| Sequencing System	No. of Exomes per Run at 50×	No. of Exomes per Run at 100×
MiSeq Series	1	N/A
NextSeq Series		
Mid-Output Flow Cell	3	2
High-Output Flow Cell	12	6
HiSeq Series		
HiSeq 2500 System Rapid-Run Mode (Dual Flow Cell)	24	12
HiSeq 2500 System High-Output Mode (Dual Flow Cell)	156	78
HiSeq 3000 System	96	48
HiSeq 4000 System (Dual Flow Cell)	192	96

a. Estimated number of exomes sequenced per run is calculated with a mean coverage of 50× and 100×, respectively. Illumina recommends a 2 × 75 bp read length on all sequencers when using TruSeq DNA Exome.

Table 2: Exome Content with TruSeq DNA Exome and Nextera™ Exome

Coverage Specification	TruSeq DNA Exome or Nextera Exome
Target Size	45 Mb
No. of Target Exons	214,405
Target Content	coding exons
Percent of Exome Covered (by Database)	
RefSeq[a]	99.45%
CCDS[b]	98.83%
ENSEMBL[c]	99.68%
GENCODE v19[d]	99.68%

a. RefSeq - NCBI Reference Sequence Database. www.ncbi.nlm.nih.gov/refseq/. Accessed February 11, 2015.
b. CCDS - Consensus CDS (CCDS) Database. www.ncbi.nlm.nih.gov/projects/CCDS/CcdsBrowse.cgi. Accessed February 11, 2015.
c. ENSEMBL - Ensembl Genome Browser. www.ensembl.org/index.html. Accessed February 11, 2015.
d. GENCODE - GENCODE Project: Encyclopedia of genes and gene variants. www.gencodegenes.org/. Accessed February 11, 2015.

Figure 1: On-Target Enrichment—TruSeq DNA Exome delivers ≥ 80% of on-target sequencing reads for efficient, cost-effective exome sequencing.
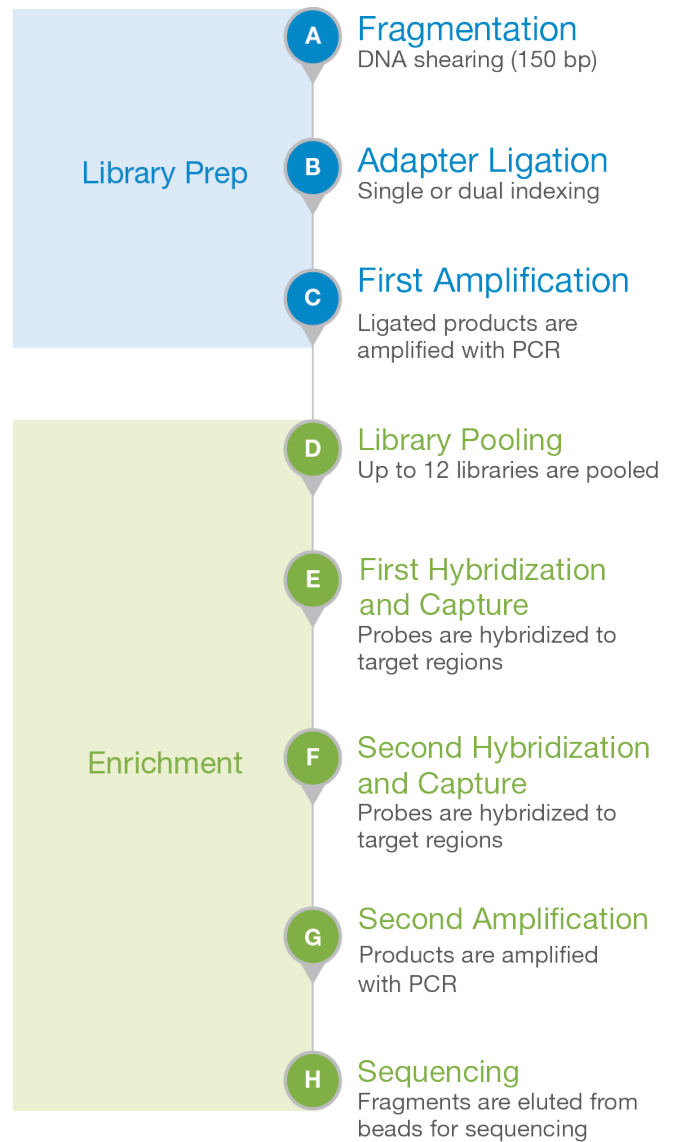
Efficient Exome Sequencing

TruSeq DNA Exome supports 12-plex pooling, enabling researchers to maximize sequencing throughput and identify variants in less time by sequencing up to 12 libraries per flow cell lane. TruSeq DNA Exome delivers ≥ 80% of on-target sequencing reads (Figure 1) and good coverage uniformity for high-confidence results. It also enables sequencing of more exomes per run, allowing researchers to maximize their budgets.

Efficient Enrichment Chemistry

TruSeq DNA Exome is optimized to deliver consistent performance for many different sample types. The library preparation workflow (Figure 2) begins with mechanical fragmentation, generating uniform fragment sizes for maximum reproducibility between libraries (Figure 2A). This mechanical shearing by Covaris ultrasonification or a similar method supports use with compromised samples, such as those containing short DNA fragments. Blunt-end DNA fragments are generated using a combination of fill-in reactions and exonuclease activity, followed by size selection with provided AMPure solid phase reversible immobilization (SPRI) beads (Beckman Coulter). Adapters containing the full complement of sequencing primer hybridization sites for single, paired-end, and indexed reads are ligated to the fragments (Figure 2B). The ligated products are amplified using PCR (Figure 2C).

Next, libraries are pooled and denatured (Figure 2D). Biotinylated probes are hybridized to the targeted regions (Figure 2E and 2F), which are enriched using streptavidin beads. After another PCR reaction (Figure 2G), fragments are eluted from the beads and ready for sequencing (Figure 2H). This streamlined workflow generates a target insert size of ~150 bp and can be completed in < 2.5 days. TruSeq DNA Exome enriches exome content for sequencing efficiently, delivering high coverage uniformity with > 85% of bases covered at 10× depth (Figure 3).

Library Prep

A — **Fragmentation**
DNA shearing (150 bp)

B — **Adapter Ligation**
Single or dual indexing

C — **First Amplification**
Ligated products are amplified with PCR

Enrichment

D — **Library Pooling**
Up to 12 libraries are pooled

E — **First Hybridization and Capture**
Probes are hybridized to target regions

F — **Second Hybridization and Capture**
Probes are hybridized to target regions

G — **Second Amplification**
Products are amplified with PCR

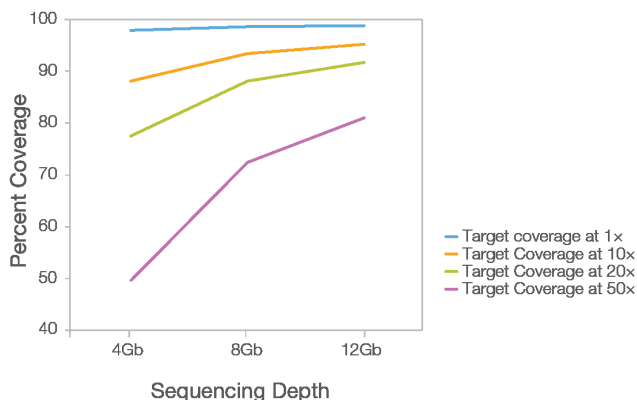H — **Sequencing**
Fragments are eluted from beads for sequencing

Figure 2: TruSeq DNA Exome Workflow—TruSeq DNA Exome combines library preparation with exome enrichment and can be completed in less than 2.5 days..
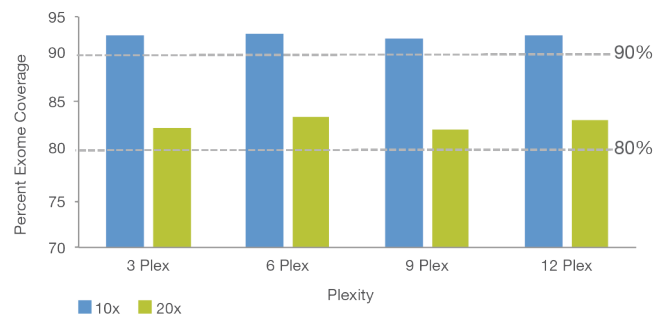

Figure 3: High Coverage Uniformity—TruSeq DNA Exome delivers uniform coverage, with > 85% of bases covered at 10× depth.

Accurate, Reliable Findings

TruSeq DNA Exome delivers exceptional target coverage over a broad range of read depths (Figure 4). Together, the reproducibility of TruSeq DNA Exome, high coverage uniformity, and SBS chemistry result in highly accurate variant calls. Over 99.65% of variant calls made following TruSeq DNA Exome and Illumina sequencing match standard reference data in the National Institute of Standards and Technology (NIST) database (Figure 5).[4,5]

Figure 4: Coverage Efficiency at Varying Depths—TruSeq DNA Exome delivers exceptional coverage across varying sequencing depth, with > 80% of targets covered up to 20× depth.
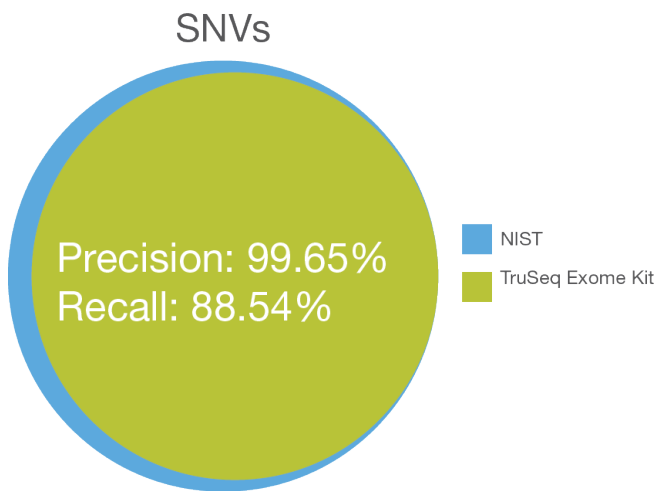

Figure 5: High Correlation with NIST Database—Variant calls made with TruSeq DNA Exome demonstrate high concordance with standard reference data. The Centre de'Etude du Polymorphism Humain (CEPH) DNA sample NA12878 was sequenced to 100× coverage depth. Single nucleotide variant (SNV) calls are reported. **Precision** is defined as the probability that a called variant is accurate. **Recall** is defined as the probability of calling a validated variant.
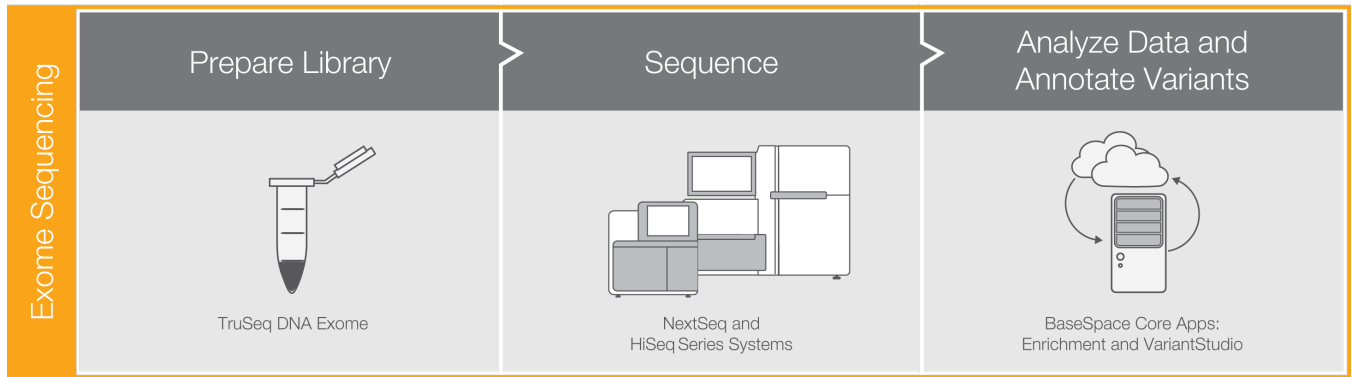
Integrated Sequencing Workflow

TruSeq DNA Exome is part of a cohesive, supported solution that guides researchers from library preparation through data analysis (Figure 7). The kit combines library preparation and exome enrichment, eliminating the need to purchase indexes, sample purification beads, or other ancillary materials. All components of TruSeq DNA Exome are designed, optimized, and analytically validated together, eliminating the need to evaluate multiple, disparate components. Expert Illumina scientists provide a single source of technical and field support for every stage of the workflow. By joining the Illumina community, researchers can harness the expertise of the Illumina support team and collaborate with the large network of scientists using Illumina technology.

Sequencing data are transferred automatically from Illumina systems to BaseSpace® Sequence Hub, the Illumina genomics computing environment. BaseSpace Sequence Hub removes much of the complexity from the typical analysis workflow, simplifying data analysis and biological interpretation. BaseSpace Sequence Hub offers an established ecosystem of integrated data analysis tools designed for biologists. With BaseSpace Apps, expert-preferred analysis tools are packaged in an intuitive, user-friendly interface, so that any researcher can access trusted analysis pipelines without previous bioinformatics experience (Figure 6). Researchers can choose to analyze exome data using the BWA Enrichment App, which uses the industry-standard BWA/GATK method, or the Isaac™ Enrichment App, which uses the fast and accurate Illumina pipeline.[6]

For biologists investigating the genetic basis of disease, the VariantStudio App enables identification and functional interpretation of disease-associated single nucleotide variants (SNVs) and insertions and deletions (indels). Researchers can rapidly filter and isolate consequential variants to enrich sequencing data with biological context. Significant findings are exported in concise reports. The VariantStudio App enables researchers to explore biological significance in a few simple steps.
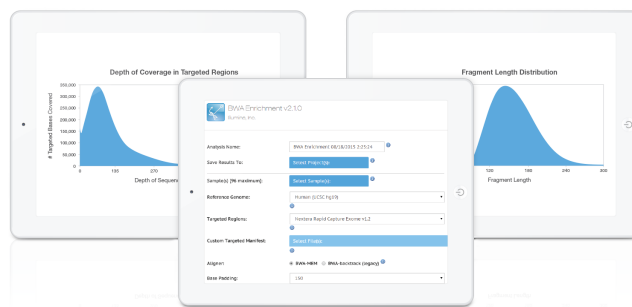
Figure 6: Simplified Data Analysis with BaseSpace Apps—TruSeq DNA Exome sequencing data can be easily and securely uploaded to BaseSpace Sequence Hub and analyzed with the BWA Enrichment App. Results are provided in easy-to-read formats.

Figure 7: Exome Sequencing Workflow—TruSeq DNA Exome is part of an integrated exome sequencing workflow that includes library preparation, sequencing, and data analysis.

Exome Sequencing Performance Comparison

Illumina offers two integrated workflow solutions for exome sequencing. Workflows are also available that combine Illumina library prep with TruSeq DNA Exome or Nextera DNA Exome, followed by exome enrichment using xGen® Universal Blockers, xGen Lockdown Reagents, and xGen Exome Research Panel v1.0, available from IDT (Table 3).

Table 3: Exome Workflow Performance Comparison

Metric	TruSeq-xGen[a]	Nextera-xGen[a]	TruSeq Exome	Nextera Exome
DNA Input	100 ng	50 ng	100 ng	50 ng
Sample Types	DNA	DNA	DNA and FFPE	DNA
Hands-On Time	5 hours	2 hours	6 hours	3 hours
Total Assay Time	2.5 days	2 days	2.5 days	2 days
Hybridization Time	4 hours	4 hours	16 hours	2 hours
On-Target %	> 91%	> 92%	> 80%	> 75%
% Coverage at 20×[b]	> 95%	> 85%	> 90%	> 85%

a. Specifications for Illumina-IDT exome enrichment workflows are based on preliminary data posted on BaseSpace Sequence Hub.
b. Percent coverage at 20× was determined for TruSeq-xGen and Nextera-xGen kits with 3.5 Gb of sequencing. Percent coverage at 20× was determined for TruSeq DNA Exome and Nextera DNA Exome with 8 Gb of sequencing.

Summary

TruSeq DNA Exome offers a streamlined, cost-effective method for identifying and understanding coding variants with exceptional data accuracy. Integration within a complete workflow consisting of leading sequencing technology and easy-to-use analysis tools enables researchers to access a single source for all their exome sequencing needs.

Learn More

To learn more about exome sequencing, visit www.illumina.com/techniques/sequencing/dna-sequencing/targeted-resequencing/exome-sequencing.html.

Ordering Information

Product	Catalog No.
TruSeq Exome Kit (24 samples)	20020614
TruSeq Exome Kit (96 samples)	20020615
IDT for Illumina – TruSeq DNA UD Indexes (24 indexes, 96 samples)	20020590
IDT for Illumina – TruSeq DNA UD Indexes (96 indexes, 96 samples)	20022370

References

1. Litchfield K, Summersgill B, Yost S, et al. Whole-exome sequencing reveals the mutational spectrum of testicular germ cell tumours. *Nat Commun*. 2015;6:5973.

2. Srivastava S, Cohen JS, Vernon H, et al. Clinical whole exome sequencing in child neurology practice. *Ann Neurol*. 2014;76:473–483.

3. Worthey EA, Mayer AN, Syverson GD, et al. Making a definitive diagnosis: successful clinical application of whole exome sequencing in a child with intractable inflammatory bowel disease. *Genet Med*. 2011;13:255–262.

4. Standard Reference Data (www.nist.gov/srd). Accessed 11 February 2015.

5. Genome in a Bottle Consortium | Advances in Biological and Medical Measurement Science (sites.stanford.edu/abms/giab). Accessed 20 February 2015.

6. Raczy C, Petrovski R, Saunders CT, et al. Isaac: ultrafast whole-genome secondary analysis on Illumina sequencing platforms. *Bioinformatics*. 2013;29:2041–2043.

For Research Use Only. Not for use in diagnostic procedures.