# NextSeq® Series Exome Sequencing Solution

A cost-effective, high-coverage exome sequencing solution that delivers accurate variant calling.

### Highlights

- **Fast Sample-to-Data Exome Solution**
  Simple workflow with the lowest hands-on time
- **Comprehensive Exome Coverage**
  Interrogate more of the exome than ever before, even in challenging regions
- **Easy Data Analysis**
  Walk-away, sample-to-results solution with analysis performed locally or in the cloud
- **Most Accurate Variant Detection**
  Accurate calls and low detection limit to identify common mutations and rare somatic events
- **End-to-End Illumina Support**
  Illumina scientists and engineers are there every step of the way, providing installation, training, applications, and data analysis support

## Introduction

The NextSeq Series Exome Sequencing Solution enables researchers to investigate the protein-coding (exonic) regions of the genome, identifying variants for population genetics, genetic disease, and cancer studies efficiently and cost-effectively. It leverages industry-leading Illumina next-generation sequencing (NGS) technology responsible for > 90% of global exome sequencing, delivering the best data quality and highest accuracy to identify true coding variants. The NextSeq Series Exome Sequencing Solution includes integrated library preparation and exome enrichment, push-button sequencing, and simple data analysis. With minimal hands-on time, the NextSeq Series Exome Sequencing Solution is the most flexible, comprehensive tool for interrogating more of the exome quickly and efficiently.


**Figure 1: NextSeq Series Sample-to-Answer Exome Sequencing Workflow**—The simple NextSeq workflow delivers highly accurate sequencing data. Data analysis includes alignment and variant calling.

## A Simple and Efficient Exome Sequencing Workflow

The NextSeq Series exome sequencing workflow simplifies exome sequencing, enabling researchers to maximize their productivity (Figure 1). It begins with library preparation and exome enrichment using the TruSeq® Exome or TruSeq Rapid Exome Kits. Dual sequencing output modes allow researchers to scale their exome studies based on sample and output needs.

Base calls are generated on the instrument (Figure 2) and data analysis (including alignment and variant calling) can be performed in BaseSpace®, the Illumina genomics computing environment. With its intuitive interface accessed through a common web browser, BaseSpace provides researchers with access to a rich ecosystem of commercial and open-source DNA data analysis software tools designed primarily to analyze Illumina data.

**Figure 2: NextSeq System**—The NextSeq Series of sequencing systems harnesses the latest advances in SBS chemistry and the simplest workflow in the industry.

The NextSeq Series also offers cross-application flexibility, enabling researchers to transition easily between sequencing projects (Figure 3). The system is fully compatible with a wide range of library preparation kits from Illumina and third parties, enabling an easy transition between exome, RNA-Seq, and whole-genome sequencing (WGS). For example, researchers can pair exome sequencing with transcriptome sequencing (RNA-Seq) to assess whether identified variants alter transcript expression or expand their studies to explore areas beyond just the coding regions with WGS. A wide range of customizable Illumina targeted resequencing solutions are also available to validate variants discovered from any sequencing application. With the NextSeq 550 System, researchers can perform NGS and array scanning on the same platform for further exploration or confirmation of copy number variants detected through sequencing.

**What is a "true coding variant"?**

A true coding variant is an accurate base call that differs from the consensus sequence within a coding region. It is not a false positive (where a variant is called but does not truly exist) or a false negative (where a variant that truly exists is not called). A system with a high false positive call rate requires extensive downstream validation, increasing costs and experimental time. A system with a high false negative call rate is failing to detect potentially important findings, often in regions that are highly repetitive or that contain homopolymer stretches.

The NextSeq Series leverages Illumina industry-leading sequencing technology to deliver the highest number of true coding variants.

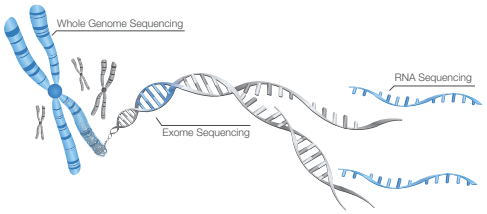


**Figure 3: NextSeq Series Sequencing Applications**—The flexible NextSeq Series enables researchers to transition seamlessly between sequencing applications to advance their research.

## Integrated Library Preparation and Exome Enrichment

Researchers can choose between 2 kits with different chemistries to initiate their studies: TruSeq Exome and TruSeq Rapid Exome Kits.* Both kits integrate library preparation and exome enrichment, enabling researchers to rapidly identify true coding variants within 45 Mb of expertly selected exonic content (Table 1).

For the most efficient exome, the TruSeq Rapid Exome Kit achieves library preparation with a tagmentation enzyme; enrichment of 1-96 exomes can be completed in 1 days, with < 3 hours of hands-on time. With a low input amount of 50 ng, the TruSeq Rapid Exome Kit enables access to precious DNA samples, while maintaining high coverage uniformity and enrichment rates. High sensitivity enables the detection of low frequency variants for accurate identification of germline and rare somatic mutations.

When maximum sensitivity and uniformity are required, the TruSeq Exome Kits delivers ≥ 80% of on-target sequencing reads with high coverage uniformity at > 85% of bases covered at 10× depth. The TruSeq Exome Kit workflow begins with 100 ng input DNA that is mechanically fragmented, generating uniform fragment sizes for maximum reproducibility between libraries. The TruSeq Exome chemistry is optimized for many different sample types including formalin-fixed paraffin embedded (FFPE) tissue.

## Delivers True Variant Calls

Obtaining true variant calls is as much a function of sequencing accuracy as it is of high-quality library preparation and enrichment. With Illumina NGS technology, the NextSeq Series delivers industry-leading sequencing accuracy of > 80% of sequenced bases over Q30† at 2 × 75 bp. It can successfully sequence even the most difficult regions (GC-rich, homopolymers), yielding a higher percentage of true coding variants than other high-throughput desktop sequencers. Its low false positive and false negative rates drastically reduce the time and cost of downstream validation. New NextSeq v2 reagent kits are optimized to improve base calling and data quality even further. By offering the highest data quality, NextSeq Systems offer the best option for capturing the full utility of the exome.

The easy NextSeq workflow simplifies exome sequencing. It takes less than 10 minutes to load and initiate the system. Sequencing is completed in as little as 18 hours for up to 15 samples using the highest output mode and paired-end 75 bp read lengths. The flexible, scalable dual sequencing output modes, multiple reagent kit configurations (from 3 to 12 pooling levels), and ability to handle a range of sample sizes enable researchers to tune and optimize their exome studies easily.

Based on industry-leading Illumina sequencing by synthesis chemistry, the NextSeq Series enables researchers to compare and integrate the data it generates with data from studies performed on other Illumina systems. For example, NextSeq Series exome sequencing data can be integrated with data from follow-up studies performed with more targeted panels on the Illumina MiSeq® System or large-scale exome sequencing studies run on an Illumina HiSeq® System (Table 2).

**Table 1: TruSeq Exome and TruSeq Rapid Exome Kits Provide Comprehensive Exome Coverage**

| Application | TruSeq Exome | TruSeq Rapid Exome |
|---|---|---|
| **Coverage Specifications** | | |
| Probe Target Size | 45 Mb | 45 Mb |
| Number of Target Exons | 214,405 | 214,405 |
| **Percent of Exome Covered by Database** | | |
| RefSeq | 98.3% | 98.3% |
| CCDS | 98.6% | 98.6% |
| Ensembl | 97.8% | 97.8% |
| GENCODE v12 | 98.1% | 98.1% |
| **Kit Performance** | | |
| On-Target Sequencing Reads | ≥ 80% | 75% |
| Coverage at 20× with 8 Gb | 85% | 80% |
| Total Assay Time | 2.5 days | 1 day |
| Sample Types | DNA and FFPE | DNA |
| DNA Input | 100 ng | 50 ng |

* The open-platform NextSeq Series can also accommodate library preparation and enrichment kits developed by third-party providers.

† Q30 = 1 error in 1,000 base calls or an accuracy of 99.9%



**Table 2: Illumina Sequencing Systems Exome Throughput**

| Sequencing System | Run Mode | Flow Cell | No. of Exomes/ Run[a] |
|---|---|---|---|
| MiSeq Series | N/A | Single | 1 |
| NextSeq Series | Mid-Output | Single | 3 |
| | High-Output | Single | 15 |
| HiSeq 2500 | Rapid-Run | Dual | 24 |
| | High-Output | Dual | 156 |
| HiSeq 3000 | N/A | Single | 96 |
| HiSeq 4000 | N/A | Dual | 192 |

a. Exome calculations performed with 2 × 75 bp read lengths. Average depth per exome was 50× (4 Gb).

## Simplified Bioinformatics in BaseSpace

The Illumina BaseSpace environment removes much of the complexity from the typical informatics workflow. Bases generated on a NextSeq System are instantly and securely transferred, analyzed, and stored in BaseSpace (Cloud or Onsite), delivering annotated variants in < 5 hours per sample. Analytic tools from NextBio® (an Illumina company) can be used to annotate and filter variants, as well as integrate and interpret genomic data in the context of other molecular and phenotypic data. Researchers can choose to analyze DNA data using the industry-standard BWA/GATK method or the fast and accurate Illumina Isaac pipeline.[1] Analysis results, including coverage statistics and annotated SNPs and indels are presented in intuitive, easy-to-interpret reports (Figure 4).

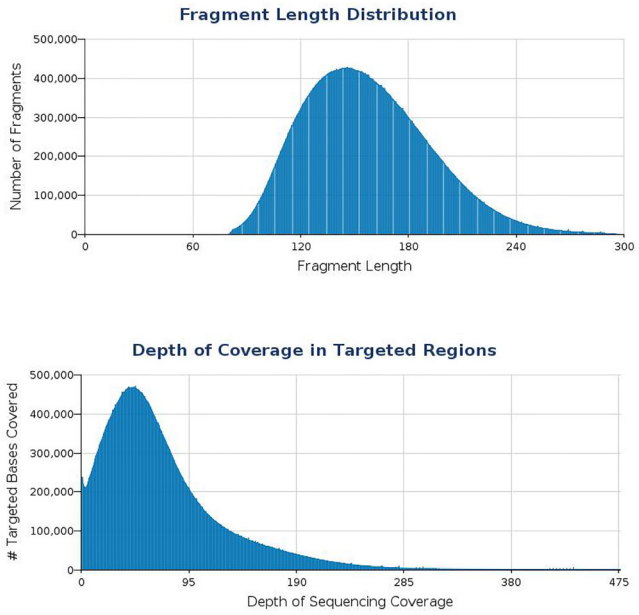


**Figure 4: Storage and Analysis of NextSeq Series Data in the BaseSpace Cloud**—NextSeq Series data can be securely and seamlessly uploaded to the BaseSpace cloud for fast, cost-effective analysis and storage.

The BaseSpace environment also includes BaseSpace apps, a growing community of software solutions for visualization, analysis, and sharing. Because Illumina NGS technology is the most established and broadly adopted sequencing solution, researchers can take advantage of a large collection of commercial and open-source sequencing analysis software designed primarily for processing Illumina data. Through its intuitive user interface and this rich ecosystem of third-party tools and apps, BaseSpace enables researchers to identify known disease variants quickly.

## Ordering Information

| System Name | Catalog No. |
|---|---|
| NextSeq 500 System | SY-415-1001 |
| NextSeq 550 System | SY-415-1002 |
| **Output Kit Name** | **Catalog No.** |
| NextSeq 500 Mid-Output v2 Kit (150 cycles) | FC-404-2001 |
| NextSeq 500 High-Output v2 Kit (150 cycles) | FC-404-2002 |
| NextSeq 500 Mid-Output v2 Kit (300 cycles) | FC-404-2003 |
| NextSeq 500 High-Output v2 Kit (300 cycles) | FC-404-2004 |
| NextSeq 500 High-Output v2 Kit (75 cycles) | FC-404-2005 |
| **Library Preparation Kit Name** | **Catalog No.** |
| TruSeq Exome Library Prep Kit (24 samples) 8 reactions × 3 plex | FC-150-1001 |
| TruSeq Exome Library Prep Kit (48 samples) 8 reactions × 6 plex | FC-150-1002 |
| TruSeq Exome Library Prep Kit (72 samples) 8 reactions × 9 plex | FC-150-1003 |
| TruSeq Exome Library Prep Kit (96 samples) 8 reactions × 12 plex | FC-150-1004 |
| TruSeq Rapid Exome Library Prep Kit (8 samples) 8 reactions × 1 plex | FC-144-1000 |
| TruSeq Rapid Exome Library Prep Kit (24 samples) 8 reactions × 3 plex | FC-144-1001 |
| TruSeq Rapid Exome Library Prep Kit (48 samples) 8 reactions × 6 plex | FC-144-1002 |
| TruSeq Rapid Exome Library Prep Kit (72 samples) 8 reactions × 9 plex | FC-144-1003 |
| TruSeq Rapid Exome Library Prep Kit (96 samples) 8 reactions × 12 plex | FC-144-1004 |

**Table 3: NextSeq Series Performance Parameters**

| Flow Cell Configuration | Read Length (bp) | Output (Gb) | Run Time | Data Quality | Required DNA Input |
|---|---|---|---|---|---|
| High-Output Flow Cell<br>Up to 400 M single reads<br>Up to 800 M paired-end reads | 2 × 150 | 100–120 | 29 hours | 80% > Q30 at 2 × 75 bp | 50 ng with TruSeq Rapid Exome<br>100 ng with TruSeq Exome |
| | 2 × 75 | 50–60 | 18 hours | | |
| | 1 × 75 | 25–30 | 11 hours | | |
| Mid-Output Flow Cell<br>Up to 130 M single reads<br>Up to 260 M paired-end reads | 2 × 150 | 32.5–39 | 26 hours | | |
| | 2 × 75 | 16.25–19.5 | 15 hours | | |

Total times include cluster generation, sequencing, and base calling on a NextSeq System. Install specifications are based on Illumina PhiX control library at supported cluster densities (between 129 and 165 K/mm² clusters passing filter). Actual performance parameters can vary based on sample type, sample quality, and clusters passing filter. The percentage of bases > Q30 is averaged over the entire run.

### NextSeq Series System Specifications

| |
|---|
| **Instrument Configuration** |
| RFID tracking for consumables |
| **Instrument Control Computer (Internal)[a]** |
| Base Unit: Dual Intel Xeon ES-2448L 1.8 GHz CPU |
| Memory: 96 GB RAM |
| Hard Drive: 750 GB |
| Operating System: Windows 7 embedded standard |
| **Operating Environment** |
| Temperature: 19°C to 25°C (22°C ± 3°C) |
| Humidity: Noncondensing 20%–80% relative humidity |
| Altitude: Less than 2,000 m (6,500 ft) |
| Air Quality: Pollution degree rating of II |
| Ventilation: Up to 2,048 BTU/hr @ 600 W |
| For Indoor Use Only |
| **Light Emitting Diode (LED)** |
| 520 nm, 650 nm; Laser diode: 780 nm, Class IIIb |
| **Dimensions** |
| W×D×H: 53.3 cm × 63.5 cm × 58.4 cm (21.0 in × 25.0 in × 23.0 in) |
| Weight: 83 kg (183 lbs) |
| Crated Weight: 151.5 kg (334 lbs) |
| **Power Requirements** |
| 100–120 VAC 15 A |
| 220–240 VAC 10 A |
| **Radio Frequency Identifier (RFID)** |
| Frequency: 13.56 MHz |
| Power: Supply current 120 mA, RF output power 200 mW |
| **Product Safety and Compliance** |
| NRTL certified IEC 61010-1 |
| CE marked |
| FCC/IC approved |

a. Computer specifications are subject to change.

## Summary

Delivering the most accurate data, the NextSeq Series Exome Sequencing Solution offers researchers the simplest and most reliable exome sequencing method to identify true coding variants. The streamlined workflow, comprehensive exome coverage, and flexible sample sizes enable researchers to efficiently and accurately interrogate more of the exome. Rapid library preparation and enrichment, industry-leading sequencing accuracy, and simple data analysis enable researchers to call the correct exome variants and move their research forward faster.

## Learn More

To learn more about the next revolution in exome sequencing, visit: www.illumina.com/applications/sequencing/dna_sequencing/exome-sequencing.html

## Reference

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

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

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


Pub. No. 770-2013-047 Current as of 08 December 2015

illumina®