Benefits of NGS
Targeted Resequencing
Taking science further: beyond qPCR and Sanger sequencing for somatic and germline variant detection
# Table of Contents

Why is Now the Time to Consider NGS Targeted Resequencing?  
- The Dramatic Decrease in the Cost of Next-Generation Sequencing  
- The Focused Power of NGS Targeted Resequencing  
- Grant Funding Trends for NGS Targeted Resequencing

When Does Targeted Resequencing Make the Most Sense?

What are the Benefits of Targeted Resequencing vs PCR and Sanger Sequencing?  
- Targeted Resequencing Allows Deep Sequencing and Higher Sensitivity  
- Targeted Resequencing Delivers Faster Turnaround for Clinical Research  
- Targeted Resequencing Provides Higher Discovery Power and Mutation Resolution  
- Next-Generation Sequencing Enables Integrated Genomics  
- Summary of Targeted Resequencing Benefits

Illumina NGS Targeted Resequencing Solutions  
- Targeted Resequencing Methods  
- Illumina Targeted Resequencing Workflows  
- Library Preparation  
- Sequencing Chemistry and Sequencing Systems  
- Simplified Data Analysis

Summary

Glossary

References

This document highlights recent publications that demonstrate the use of Illumina technologies in immunology research. To learn more about the platforms and assays cited, visit [www.illumina.com](http://www.illumina.com).
Why is Now the Time to Consider NGS Targeted Resequencing?

The Dramatic Decrease in the Cost of Next-Generation Sequencing

Over the past decade, molecular biologists and translational researchers have focused on the identification of somatic or germline DNA variants with strong associations to cancer, neurobiological disorders, and other complex diseases. Quantitative PCR, Sanger sequencing (capillary electrophoresis sequencing), and microarray technology have all played an important role in these genetic studies. During the same time period, the massively parallel sequencing technology known as next-generation sequencing (NGS) has revolutionized the biological sciences. From the emergence of the Illumina Genome Analyzer in 2006 to the release of the HiSeq X® Ten in 2014, the data output of NGS has outpaced Moore’s law—more than doubling each year (Figure 1). The rapid drop in sequencing cost and the massive increase in data output, has resulted in NGS fundamentally changing the kinds of questions scientists can ask and answer. Researchers can now analyze the complete human genome in a single sequencing experiment, and sequence thousands to tens of thousands of genomes in a single year (Figure 2). As Eric Lander, founding director of the Broad Institute of MIT and Harvard stated, “As costs continue to come down, we are entering a period where we are going to be able to get the complete catalog of disease genes. This will allow us to look at thousands of people and see the differences among them, to discover critical genes that cause cancer, autism, heart disease, or schizophrenia.”

While the drop in cost for whole-genome sequencing (WGS) is an exciting development for science, the reduction in cost and the increasing simplicity of other NGS methods, such as targeted resequencing, have made the benefits of NGS accessible to the wider research community. With a growth rate on the rise, targeted resequencing is proving to be a powerful tool for somatic and germline variant detection.

See how genomics and targeted resequencing are empowering researchers around the globe: Global perspectives on the impact of next-generation sequencing.
The Focused Power of NGS Targeted Resequencing

With targeted resequencing, a subset of genes or target regions (a sequencing panel) are isolated and enriched before sequencing, efficiently and cost-effectively focusing the power of NGS. Targeted resequencing offers several significant advantages. It enables deep sequencing (sequencing at much higher coverage levels), which allows greater confidence over Sanger sequencing for calling variants or low-frequency alleles in a given region of interest.²,³ When speed is critical to success, targeted resequencing can also provide fast turnaround times, due to a higher multiplexing capacity, lower data analysis requirements, and the ability to sequence tens to thousands of targets in a single experiment.⁴

Targeted resequencing can reveal variants, such as low-frequency variants that would be more expensive or more challenging to identify with PCR or Sanger sequencing.⁵,⁶ The ability to detect low-frequency variants can enable identification of novel functional variants, facilitate biomarker discovery, or lead to the identification of clinically relevant targets for translational research.⁴,⁵ Targeted resequencing is particularly useful for the discovery of somatic mutations in complex samples such as cancerous tumors mixed with germline DNA.⁷–⁹ Whether performing cancer studies, microbial genomics, agrigenomics, or molecular epidemiology, researchers can target regions of the genome relevant to their specific interests.

![Figure 2: Targeted Resequencing Data Output vs Sanger Sequencing](image)

For $150 USD per sample, targeted resequencing can produce nearly 12× more data and cover hundreds more targets than Sanger sequencing. Calculations assume an NGS targeted resequencing run using the TruSeq® Custom Amplicon Low Input Library Prep Kit, MiSeq® v2 Sequencing Reagent Kit, a read length of 450 bp, 700 targets, and 50× coverage. Sanger sequencing assumptions include a read length of 800 bp, 30 targets, and a price per target (ie, price/reaction) of $5 USD.
Grant Funding Trends for NGS Targeted Resequencing

With the declining cost of sequencing overall, and improvements in targeted resequencing methods, targeted resequencing is accelerating the pace of research and driving high impact publications. From 2005 to 2015, the number of publications with targeted resequencing has grown from a handful of studies to over 300 publications (Figure 3). As the impact of targeted resequencing continues to grow, grant funding trends for targeted resequencing are also on the rise (Figure 4). Data from the National Institutes of Health (NIH) show that funding for research including NGS targeted resequencing has steadily increased each year and is 34% higher than in 2012. These trends show that studies using NGS targeted resequencing methods are benefiting from increased funding and resulting in more publications every year.

Figure 3: Publications with Targeted Resequencing are Increasing—In PubMed, the number of times in which the keywords for “TGRS” have been mentioned in titles and abstracts of papers has grown from ~10 to > 300 in the last 10 years. Keywords include: “targeted resequencing” OR “targeted resequencing” OR “NGS TGRS” OR “TGRS” OR “amplicon sequencing” OR “enrichment sequencing.”

Figure 4: Grant Funding for Targeted Resequencing on the Rise—NIH funding for studies with targeted resequencing has seen a 34% growth rate between 2012 and 2015.
When Does Targeted Resequencing Make the Most Sense?

With the rapid pace of changing technology, deciding when to add a new technique to the repertoire, or whether to make a complete transition from one method to another can be difficult to assess. Deciding when to use PCR, Sanger sequencing, targeted resequencing, or other NGS approaches such as exome sequencing or WGS depends on a combination of factors. These include the number of samples, the total amount of sequence in the target regions, budgetary considerations, and ultimately, the overall goals of the research study. Sanger sequencing and PCR are typically good choices when the number of target regions is low (1–20 target regions), and when the study aims are limited to screening or identification of known targets or variants (Figure 5). At the opposite end of the sequencing spectrum, exome sequencing and WGS are excellent methods for comprehensive genetic analysis and variant discovery. Targeted NGS offers a balanced choice between these options that supports both screening and variant discovery study designs. It is the most cost-effective approach for the sequencing of tens to thousands of genes with a high number of samples (Figure 6). With the ability to sequence multiple genes across multiple samples simultaneously, targeted resequencing methods save time and resources compared to traditional iterative methods. Moreover, targeted NGS methods preserve precious sample material by requiring lower DNA input—an important consideration for applications such as analysis of forensic DNA samples or cancer biopsy samples (Figure 6).

Figure 5: Options for WGS, Exome Sequencing, Targeted Resequencing, Sanger Sequencing, and PCR—PCR and Sanger sequencing are effective approaches for variant screening studies when the total number of samples is low. For variant screening studies where the sample number is high, NGS amplicon sequencing is more efficient and cost-effective. For discovery-related applications, any NGS approach will provide higher discovery power compared to PCR/Sanger sequencing. The discovery power will increase as the total target sequence increases. Targeted resequencing (amplicon or enrichment method) is the most cost-effective solution when sequencing more than 1–20 target regions.

Figure 6: Cost Effectiveness for Targeted Resequencing vs PCR and Sanger Sequencing—The area above the line represents higher cost-effectiveness with targeted DNA sequencing compared to Sanger sequencing or qPCR.
What are the Benefits of Targeted Resequencing vs PCR and Sanger Sequencing?

Targeted Resequencing Allows Deep Sequencing and Higher Sensitivity

As with all targeted NGS methods, the targeted resequencing approach provides a rapid and cost-effective alternative to single-gene testing. However, the most powerful aspect of targeted resequencing may be the ability to sequence a given region of interest at much higher coverage levels compared to WGS or Sanger sequencing. WGS is typically performed at 30X–75X coverage while targeted NGS enables sequencing depths of 5000X or higher (Figure 7). From a research perspective, deep sequencing translates into higher sensitivity (ability to detect low-frequency variants). Sanger sequencing can provide a limit of detection down to 20% allele frequency while NGS targeted resequencing provides a limit of detection down to 1%.\textsuperscript{13,14} Deep sequencing and the resulting high-sensitivity can be critical for certain kinds of studies, such as the detection of low-frequency subclones within heterogeneous tumors or the detection of somatic mosaicism.\textsuperscript{3,15,16} Several studies have shown that Sanger sequencing does not have the sensitivity required to detect mosaic variants consistently.\textsuperscript{3,15} Dr. Saumya Jamuar, cofounder of Global Gene Corporation, used the TruSeq Custom Amplicon Library Prep Kit on the MiSeq System to identify rare mosaic mutations that lead to neural malformations. Dr. Jamuar discussed targeted resequencing vs. Sanger sequencing for somatic variant detection:

> Actually, there weren’t many advantages to using Sanger sequencing for variant detection. It was a lengthy process because Sanger sequencing proceeds gene by gene. With NGS, all we had to do was put all our genes of interest together in one panel, one reaction, and we would get the results. Sanger sequencing has a limit of detection of about 15%. This isn’t low enough to detect somatic mosaicism.”

Dr. Saumya Jamuar, Genetics Service Consultant at KK Women and Children’s Hospital in Singapore and cofounder of Global Gene Corporation

<table>
<thead>
<tr>
<th>Whole-genome sequencing</th>
<th>75X</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exome sequencing</td>
<td>200X</td>
</tr>
<tr>
<td>Targeted sequencing</td>
<td>5,000X</td>
</tr>
</tbody>
</table>

Figure 7: Sequence Regions of Interest at Higher Resolution—With more reads focused on fewer regions, targeted resequencing achieves deeper coverage and higher resolution compared to WGS or exome sequencing.
Targeted Resequencing Delivers Faster Turnaround for Clinical Research

Clinical or translational studies require tests that are not only accurate, but also rapid and cost-effective. Tests must be accurate because they can lead to the development of divergent diagnosis, risk stratification, and drug treatments. Speed is also a critical factor when assessing clinical samples for translational research. While Sanger sequencing has been considered the gold standard in clinical research for many years, the process can be time consuming and inefficient when multiple gene tests and many samples are required (Figure 8). Furthermore, limited tissue availability can make iterative Sanger gene testing nearly impossible. Targeted resequencing is becoming more prevalent with clinical and translational researchers because it effectively addresses many of these challenges. A study from the University Hospital of Cologne Germany comparing Sanger sequencing and targeted resequencing methods found that the multiplexing capacity of targeted resequencing led to reduced turnaround times for analysis of tumor samples. In addition, the study authors found that targeted resequencing required significantly less input material and proved to be a more cost-effective approach, compared to the Sanger method.

Figure 8: Case Study: Targeted Resequencing Supports Faster Turnaround Times—A translational researcher investigating the genetics of epilepsy wants to interrogate 3 genes with a total of 94 possible variants across 10 samples (ATP1A2 = 34 targets, ATP13A2 = 47 targets, and CDKL5 = 13 targets). With 1 reaction per variant, Sanger sequencing would require a total of 20 reaction plates (10 samples x 1 plate x 2 reactions (forward and reverse) = 20 x 96-well plates. Assuming 1 full-time employee preps 4 plates per day, the Sanger method would require 5 days. In contrast, using targeted resequencing with sample multiplexing (library preparation with the TruSeq Custom Amplicon Library Prep Kit and sequencing on the MiniSeq™ System), the same study can be performed in 1.5 days. Multiplexing (combining multiple samples into a single sequencing run) combined with the ability to sequence thousands of targets simultaneously enables a faster turnaround time compared to other technologies.

Dr. John Robinson, research manager of the Blackburn Cardiovascular Genetics Laboratory at the Robarts Research Institute, uses a custom targeted resequencing panel (Nextera® Rapid Capture Enrichment) on the MiSeq® System to profile genetic variants associated with dyslipidemias and related metabolic disorders. Dr. Robinson discussed his experience with targeted resequencing vs. Sanger sequencing for the identification of variants associated with lipid metabolism:

We found that an individual subject might take a month to Sanger sequence before we found the causative mutation, and that’s with all the Sanger costs and the costs for the labor, PCR, etc. With targeted resequencing, within a 2-week period we get 24, 700 kb sequencing data files that have been annotated and categorized on rareness and pathogenicity. That’s 24 people every 2 weeks. So the timeframe is just impossible to compare to a Sanger 1000-base-pair-at-a-time sequencing process.”

Dr. John Robinson, research manager of the Blackburn Cardiovascular Genetics Laboratory at the Robarts Research Institute
Targeted Resequencing Provides Higher Discovery Power and Mutation Resolution

While PCR and Sanger sequencing offer quick, familiar workflows, targeted resequencing has higher discovery power (ability to identify novel variants), and high variant resolution (single base resolution). Depending on the research goals, PCR, Sanger sequencing, and targeted resequencing provide different levels of information. PCR can indicate whether a variant is present, but cannot provide single-base resolution. Sanger sequencing can cost-effectively identify specific variants with single-base resolution for a small number of genes or target regions. In contrast to qPCR and Sanger sequencing, targeted resequencing can identify variants across thousands of target regions, down to single-base resolution, in a single experiment (Figure 9). Furthermore, due to the larger scale of sequence interrogation and targeted sequence capture methods, the possibility of detecting novel variants is much higher.

![Figure 9: Targeted Resequencing Delivers Higher Discovery Power and Highest Mutation Resolution](image)

Next-Generation Sequencing Enables Integrated Genomics

While the detection of somatic and germline variants has provided important insights into the genetic basis of disease, many researchers are now leveraging integrated genomics analysis to obtain a multifaceted view of biological processes and disease pathogenesis. In a research study from the Netherlands Cancer Institute, comprehensive genomic, transcriptomic, and proteomic analysis was used to identify 3 genetic subtypes within a large cohort of invasive lobular carcinoma (ILC) histological samples. The subtypes demonstrated differences in mRNA profiles, hormone characterization, and significant differences in prognostic indicators. For cutaneous melanoma, data analysis suggests that integrating multiple types of genomic data leads to prognostic models with an improved prediction performance. For investigators seeking to integrate multiple genome-wide techniques, NGS technology provides a single platform for genomics, transcriptomics, and epigenetics research.
Summary of Targeted Resequencing Benefits

With readily available predesigned targeted sequencing panels, easy online tools for custom panel design, and the emergence of desktop NGS sequencers, targeted resequencing has become simpler and more accessible than ever before. The overall cost of NGS sequencing and targeted resequencing continues to fall and easy bioinformatics solutions are removing the previous challenges associated with NGS data analysis. Multiple types of library preparation kits for targeted resequencing have expanded the choices available to scientists and researchers. These include various panels with preselected content, and the ability to target specific regions of interest with online methods for designing and ordering custom sequencing panels. With significant advantages compared to PCR and Sanger sequencing, more and more biologists and translational researchers are choosing NGS targeted resequencing for their variant detection studies (Table 1).

Table 1: Comparison of PCR, Sanger Sequencing, and Targeted Resequencing

<table>
<thead>
<tr>
<th>Benefits</th>
<th>PCR</th>
<th>Sanger Sequencing</th>
<th>Targeted Resequencing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benefits</td>
<td>• High sensitivitya</td>
<td>• Cost-effective sequencing low numbers of targets (1–20 targets)</td>
<td>• Higher sequencing depth enables higher sensitivity (down to 1%)</td>
</tr>
<tr>
<td></td>
<td>• Familiar workflow</td>
<td>• Familiar workflow</td>
<td>• Higher discovery powerb</td>
</tr>
<tr>
<td></td>
<td>• Capital equipment already placed in most labs</td>
<td>• Current gold standard in sequencing</td>
<td>• Higher mutation resolutionc</td>
</tr>
<tr>
<td>Challenges</td>
<td>• Can only interrogate a limited set of mutations</td>
<td>• Low sensitivity (down to 20%)</td>
<td>• Produce more data with the same amount of input DNAd</td>
</tr>
<tr>
<td></td>
<td>• Virtually no discovery power</td>
<td>• Low discovery power</td>
<td>• Higher sample throughput with sample multiplexing</td>
</tr>
<tr>
<td></td>
<td>• Limited mutation resolution</td>
<td>• Not as cost-effective for high numbers of targets (&gt; 1–20 targets)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Low scalability due to increasing sample input requirements</td>
<td>• Low scalability due to increasing sample input requirements</td>
<td></td>
</tr>
</tbody>
</table>

a. sensitivity = allele frequency limit of detection  
b. discovery power = ability to identify novel variants 
c. mutation resolution = identification of mutations from large chromosomal rearrangements down to single-nucleotide variants. 
d. 10 ng DNA will produce ~1 kb with CE sequencing or ~300 kb with targeted resequencing (250 bp amplicon length × 1536 amplicons with TruSeq Custom Amplicon workflow)
See how researchers are performing integrated genomics analysis with NGS technology:

Molecular systems evaluation of oligomerogenic APP<sup>E693Q</sup> and fibrillogenic APP<sup>KM670/671NL/PSEN<sub>1</sub>Δ exon9</sup> mouse models identifies shared features with human Alzheimer’s brain molecular pathology.


Integrated clinical, whole-genome, and transcriptome analysis of multisampled lethal metastatic prostate cancer.


To learn how Dr. Andrew Fellowes is developing tumor profiling tests using Illumina TruSeq Custom Amplicon and the MiSeq® System, read the Illumina iCommunity article, *Sequencing to Inform Cancer Treatment*.

To read more about how Dr. Graham Taylor is harnessing the speed of targeted resequencing, read the Illumina iCommunity article, *Speed of MiSeq System Critical to Clinical Use*.
Illumina NGS Targeted Resequencing Solutions

Targeted Resequencing Methods

Illumina currently supports 2 methods for targeted resequencing: hybrid capture-based enrichment sequencing and amplicon sequencing. With enrichment sequencing, regions of interest are captured by sequence-specific hybridization probes. The second method, amplicon sequencing, involves amplification and purification of regions of interest using primer-mediated, highly multiplexed oligo sets. Enrichment sequencing enables larger gene content and more comprehensive profiling for all variant types, which requires longer turnaround time with more hands-on time. Amplicon generation is a more affordable, easier workflow than enrichment sequencing, supports smaller gene content, and is ideal for analyzing single nucleotide variants and insertions/deletions (indels). Both of these highly multiplexed approaches enable a wide range of applications for the discovery, validation, or screening of genetic variants (Figure 10).

![Figure 10: Comparison of Targeted Region Size between Amplicon and Enrichment Methods](image)

Figure 10: Comparison of Targeted Region Size between Amplicon and Enrichment Methods — Enrichment sequencing methods enable capture of much larger total content than amplicon generation.

Targeted gene sequencing panels are useful tools for analyzing specific mutations in a given sample. Focused panels contain a select set of genes or gene regions that have known or suspected associations with the disease or phenotype under study. Gene panels can be purchased with preselected content or custom designed to include genomic regions of interest (Table 2).
Illumina Targeted Resequencing Workflows

Illumina targeted resequencing workflows offer DNA-to-data solutions that include content design (for custom panels only), library preparation, sequencing, and data analysis (Figure 11).

<table>
<thead>
<tr>
<th>Design content</th>
<th>Prepare library</th>
<th>Sequence</th>
<th>Analyze data &amp; annotate variants</th>
</tr>
</thead>
<tbody>
<tr>
<td>DesignStudio™</td>
<td>Amplicon library prep: 8 hours</td>
<td>≤ 24 hours</td>
<td>≤ 24 hours</td>
</tr>
<tr>
<td>(For custom panels only)</td>
<td>Enrichment library prep: 2.5 days</td>
<td>MiniSeq™, MiSeq®, and NextSeq® Systems</td>
<td>&lt; 2 hours</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Local Run Manager BaseSpace® Sequence Hub BaseSpace Variant Interpreter (beta)</td>
<td></td>
</tr>
</tbody>
</table>

Figure 11: Targeted Resequencing Workflow — The integrated workflow begins with selection of a predesigned panel or design of custom content, followed by streamlined library preparation, sequencing, and data analysis. Targeted resequencing enables cost-effective studies for a broad range of samples and applications.

Library Preparation

Predesigned Targeted Sequencing Panels

Targeted resequencing panels are useful tools for analyzing specific variants in a given gene or region of interest. Predesigned panels contain important genes or gene regions associated with a disease or phenotype. Content is selected by key opinion leaders and from curation of the scientific literature. By far the greatest advantage of working with predesigned content is that the content has already been tested—saving significant time and cost. By focusing on a subset of the genome, these panels also minimize data analysis time and decrease storage requirements. For sample screening or variant discovery, multiple genes can be assessed across many samples in parallel, saving time and reducing costs associated with running separate, iterative assays. Predesigned panels are available for many research areas including cancer, recessive pediatric onset diseases, cardiac conditions, and more (Table 2).

Key advantages of predesigned panels:
- Sequence key genes or regions of interest to high depth (500–1000x or higher), allowing identification of low-frequency variants
- Save significant time because content has already been selected and tested
- Provide a cost-effective solution to study disease-related genes
- Deliver accurate, easy-to-interpret results, identifying variants at low allele frequencies (down to 1%)
- Enable confident identification of causative or inherited mutations in a single assay

For more information on Illumina Targeted Resequencing Panels, visit the Targeted Panels Page.
### Table 2: Illumina Targeted Resequencing Solutions

<table>
<thead>
<tr>
<th>Predesigned Targeted Sequencing Panels with Enrichment-Based Library Prep</th>
<th>Key Features/Advantages</th>
<th>Cumulative Target Region Size</th>
<th>DNA Input</th>
<th>Number of Samples per MiniSeq Run(^a)</th>
<th>Sequencing System Compatibility</th>
</tr>
</thead>
</table>
| **TruSight\(^b\) One Panel** | • Targets 4813 genes associated with known clinical phenotypes  
• 1.5 day library prep | 12 Mb | 50 ng | 3 samples/run | MiniSeq\(^b\), MiSeq\(^b\), NextSeq\(^b\), and HiSeq\(^b\) Series |
| **TruSight Cardio Panel** | • Targets 174 genes related to 17 inherited cardiac conditions  
• 1.5 day library prep | 244 kb | 50 ng | 12 samples/run | MiniSeq, MiSeq, and NextSeq Series |
| **TruSight Inherited Disease Panel** | • Targets 552 genes related to severe, recessive pediatric diseases  
• 8801 target exons | 2.25 Mb | 50 ng | 8 samples/run | MiniSeq, MiSeq, NextSeq, and HiSeq Series |
| **TruSight Cancer Panel** | • Targets 94 genes associated with a predisposition towards cancer  
• Detect variants down to 5% allele frequency | 255 kb | 50 ng | 24 samples/run | MiniSeq, MiSeq, NextSeq, and HiSeq Series |

### Predesigned Targeted Sequencing Panels with Amplicon-Based Library Prep

| **TruSight HLA v2 Sequencing Panel** | • Targets 11 loci for comprehensive HLA typing  
• Sample-to-report in less than 48 hours with less than 4 hours hands-on time | 50 kb | 400 ng | 4–48 samples/run | MiniSeq, MiSeq, and NextSeq Series |
| **TruSight Tumor 15** | • FFPE compatible  
• Targets 15 genes commonly mutated in solid tumors  
• Detect variants down to 5% allele frequency | 44 kb | 20 ng | 8 samples/run | MiSeq and MiSeq Series |
| **TruSight Myeloid Sequencing Panel** | • Targets 54 genes focused on somatic mutations in myeloid malignancies  
• Detect variants down to 5% allele frequency | 141 kb | 50 ng | 8 samples/run | MiSeq, MiSeq, and NextSeq Series |
| **TruSeq\(^c\) Amplicon Cancer Panel** | • FFPE compatible  
• Targets 48 genes with mutational hotspots in frequently mutated cancer genes | > 35 kb | 150 ng for gDNA  
250 ng for FFPE | 42 samples/run | MiSeq and NextSeq Series |

### Custom Targeted Sequencing Panels

| **Nextera\(^b\) Rapid Capture Custom** | • Enrich custom content  
• 1.5 day library prep | 0.5–15 Mb | 50 ng | 1-96 samples/ run\(^b\) | MiniSeq, MiSeq, NextSeq, and HiSeq Series |
| **TruSeq Custom Amplicon v1.5** | • FFPE compatible  
• Amplify custom content | 2–650 kb | 50 ng for gDNA  
150 ng for FFPE | 1-96 samples/ run\(^b\) | MiniSeq, MiSeq, MiSeq, and NextSeq Series |
| **TruSeq Custom Amplicon Low Input** | • FFPE compatible  
• Amplify custom content  
• Low DNA input amount | 2–650 kb | 10 ng for gDNA  
10-50 ng for FFPE | 1-96 samples/ run\(^b\) | MiniSeq, MiSeq, and NextSeq Series |

---

\(^a\) With High-Output MiniSeq Sequencing Kit  
\(^b\) Sample throughput varies with experimental design and mean coverage  
\(^c\) Input amount depends on QC results with TruSeq FFPE DNA Library Prep QC Kit

For more information on Illumina Enrichment and Amplicon Sequencing solutions, visit the Targeted Resequencing Page.
Library Prep and Array Kit Selector
To find the right library prep kit, researchers can use the Library Prep and Array Kit Selector (Figure 12). This web-based, interactive tool provides a step-by-step process to choose an application-specific library prep kit. The tool helps with kit selection based on starting material, quality of starting material, area of interest, technique, and more.

Figure 12: Library Prep and Array Kit Selector — This web-based tool helps researchers determine the best kit for their needs based on project type, starting material, and application.
Custom Panels with DesignStudio™ Software
Due to the focused nature of predesigned targeted gene panels, content may not cover all targets of interest for every researcher. Custom sequencing panels enable gene or target content to be customized based on the research interests of the investigator using DesignStudio Software (Figure 13).

Figure 13: DesignStudio Interface—Researchers can use DesignStudio Software to visualize targeted genomic regions and attempted amplicons to assess design coverage and more.

Expanded Options with Illumina Concierge Services
Illumina Concierge services offer additional design support and expanded features for Illumina custom targeted resequencing projects. For example, with Concierge Services some custom targeted resequencing kits, such as the TruSeq Custom Amplicon v1.5 Library Prep and TruSeq Custom Amplicon Low Input Library Prep Kits, can incorporate unique molecular identifiers for enhanced allelic detection and increased sensitivity. Unique molecular identifiers allow for the removal of PCR duplicates, which enables the detection and quantification of individual molecules. The TruSeq Custom Amplicon Assays are also compatible with dual-strand sequencing, which reduces false positives that can arise from deamination events during formalin fixation or from other DNA lesions. Illumina Concierge Services also offers the ability to design smaller amplicons (~100 bp), increasing compatibility with fragmented DNA, such as DNA from formalin-fixed, paraffin-embedded (FFPE) tissue. Contact an Illumina representative for access to Illumina Concierge Services.
Enrichment-Based Library Prep

Enrichment-based library prep begins with Nextera tagmentation, which converts input genomic DNA into adapter-tagged libraries without the need for mechanical shearing (Figure 14). Next, libraries are denatured and biotin-labeled probes specific to targeted regions are used for hybridization. The pool is enriched for regions of interest by adding streptavidin-coated beads that bind to the biotinylated probes. DNA fragments bound to the streptavidin-coated beads via biotinylated probes are magnetically pulled down from the solution. The enriched DNA fragments are then eluted from the beads, and a second round of hybridization and capture is performed before sequencing.

![Enrichment Workflow](image)

**Enrichment Workflow**

*Use DesignStudio to create custom oligo capture biotinylated probes*

- Biotin probes
- Hybridize biotinylated probes to targeted regions
- Streptavidin beads
- Enrichment using streptavidin beads
- Elution from beads

**Figure 14: Enrichment-Based Library Prep**— This method harnesses Nextera technology to provide a streamlined protocol that combines library preparation and enrichment steps.
**Amplicon-Based Library Prep**

Amplicon sequencing includes 2 submethods depending on the product chosen. The first method employs a hybridization-extension-ligation approach, creating a single strand template from a double-stranded DNA population that is later amplified via PCR (Figure 15). The library prep begins with hybridization using a highly multiplexed pool of oligo pairs. Each oligo pair targets a specific region of interest in the genome. Following hybridization, an extension-ligation reaction extends nucleotides across the region of interest. The resulting extension-ligation templates are PCR-amplified, incorporating library-specific indexes and sequencing primers. Final reaction products are converted to a single-stranded, adapter-ligated normalized library using a bead-based protocol. Libraries are then ready for sequencing without the need for any further processing.

The second method involves a multiplexed PCR approach, amplifying the predefined targeted regions from genomic DNA (Figure 15). This library prep uses 2 oligo pools for multiplex PCR to amplify and target regions of interest in DNA. Using index adapters, libraries are indexed and further amplified, and then combined in preparation for sequencing.

**Figure 15: Amplicon Chemistry Methods**—Amplicon sequencing uses two different chemistry methods. The extension/ligation method uses an extension/ligation reaction for the initial target amplification, while multiplex PCR uses a series of PCR reactions.
Sequencing Chemistry and Sequencing Systems

After libraries are prepared, they undergo cluster generation. Libraries are injected into a flow cell and captured by surface-bound oligos complimentary to library adapters. Each library fragment is then amplified into distinct, clonal clusters through bridge amplification. After cluster generation is complete, templates are ready for sequencing.

Illumina sequencing by synthesis (SBS) chemistry is a central part of the most successful and widely adopted NGS technology in the world.* With Illumina SBS, a fluorescently labeled reversible terminator is imaged as each dNTP is added, and then cleaved to allow incorporation of the next base. Because all 4 reversible terminator-bound dNTPs are present during each sequencing cycle, natural competition minimizes incorporation bias. The result is true base-by-base sequencing—in a massively parallel fashion. Data generation on Illumina sequencing systems has been cited in more than 34,600 peer-reviewed publications to date.

To learn more about sequencing by synthesis and view a detailed animation of SBS chemistry, visit the SBS Technology Page.

Illumina sequencing systems offer user-friendly, intuitive interfaces for easy run setup and operation at every sequencing scale (Figure 16). The MiniSeq System is ideal for amplicon, targeted RNA, small RNA, and targeted gene panel sequencing. The MiSeq System supports all these methods and enables small genome sequencing. Compared to all sequencers in the Illumina desktop portfolio, the NextSeq Series offers the highest output and maximum reads per run, enabling exome, transcriptome, and targeted resequencing applications.

*Data calculations on file. Illumina, Inc. 2015.
The Sequencing Coverage Calculator

The Sequencing Coverage Calculator is a web-based tool that helps researchers determine the reagents and sequencing runs needed to arrive at the desired coverage for a particular experiment (Figure 17). Based on the specific application, input parameters, and the sequencing systems chosen, the calculator writes a table containing the number of lanes or flow cells needed for the desired coverage. Results can be downloaded in a comma-separated values (CSV) file that can be shared or edited in Excel.

Figure 17: Sequencing Coverage Calculator—This online tool enables researchers to determine reagents and sequencing runs needed for a particular coverage, before starting an experiment.
When I began working with NGS... I was pessimistic about the bioinformatics capability. That changed after I used MiSeq Reporter. When I used MiSeq Reporter and ran the same data, my alignment increased from 50% to 90%. So the software enabled me to detect mutations that otherwise would have gone undetected not because the data was bad, but because I was not using the right tools to align.”

Dr. Saumya Jamuar, Genetics Service Consultant at KK Women and Children’s Hospital in Singapore and cofounder of Global Gene Corporation
Summary

Labs worldwide are taking advantage of NGS-based targeted resequencing to interrogate somatic and germline variants with a lower cost, faster turnaround time, and lower sample input requirements compared to Sanger sequencing and PCR-based genotyping. With targeted resequencing, researchers can focus their time, expenses, and data storage resources on the most impactful regions of the genome for their area of research. Because targeted resequencing assesses a predefined set of target regions, it enables deep sequencing and higher sensitivity for calling low-frequency variants. For rapid adoption, predesigned targeted resequencing panels featuring expert-selected content allow researchers to avoid the time and expense associated with designing and testing their own panels. Targeted resequencing is a powerful tool that is helping researchers understand the pathogenesis of complex diseases, identify disease-resistant genes in agriculturally important crops, and may one day lead to more effective treatment options for future generations.

Glossary

cluster generation: An amplification reaction that occurs on the surface of an Illumina flow cell. During flow cell manufacturing, the surface is coated with a lawn of 2 distinct oligonucleotides often referred to as “p5” and “p7.” In the first step of bridge amplification, a single-stranded sequencing library (with complementary adapter ends) is loaded into the flow cell. Individual molecules in the library bind to complementary oligos as they “flow” across the oligo lawn. Priming occurs as the opposite end of a ligated fragment bends over and “bridges” to another complementary oligo on the surface. Repeated denaturation and extension cycles (similar to PCR) results in localized amplification of single molecules into millions of unique, clonal clusters across the flow cell. This process, also known as “bridge amplification” occurs in an automated, flow cell instrument called a cBot™ or in an onboard cluster module within an NGS instrument.

clusters: A clonal grouping of template DNA bound to the surface of a flow cell. Each cluster is seeded by a single, template DNA strand and is clonally amplified through bridge amplification until the cluster has roughly 1000 copies. Each cluster on the flow cell produces a single sequencing read. For example, 10,000 clusters on the flow cell would produce 10,000 single reads and 20,000 paired-end reads.

coverage level: The average number of sequenced bases that align to each base of the reference DNA. For example, a whole genome sequenced at 30× coverage means that, on average, each base in the genome was sequenced 30 times.

deep sequencing: Sequencing to high coverage levels. For example, WGS is typically performed to 30×–75× coverage while targeted NGS enables sequencing depths of 5000× or higher.

discovery power: The ability to identify novel variants.

multiplexing: A process where a unique DNA sequence is ligated to fragments within a sequencing library for downstream, in silico sorting and identification. Indexes are typically a component of adapters or PCR primers and are ligated to the library fragments during the sequencing library preparation stage. Illumina indexes are typically between 8–12 bp. Libraries with unique indexes can be pooled together, loaded into one lane of a sequencing flow cell, and sequenced in the same run. Reads are later identified and sorted via bioinformatic software.
**mutation resolution:** The size of mutation, in base pairs, a technology is able to detect. For example, karyotyping provides a mutation resolution of 5-10 Mb, while array comparative genomic hybridization provides “higher resolution” by detecting mutations down to 50 kb. NGS techniques provide the highest possible mutation resolution because they can provide single-base pair variant detection (detect the presence of a mutation) and nucleotide identification (detect the identity of a mutation).

**read:** The process of next-generation DNA sequencing involves using sophisticated instruments to determine the sequence of a DNA or RNA sample. In general terms, a sequence “read” refers to the data string of A, T, C, and G bases corresponding to the sample DNA. With Illumina technology, millions of reads are generated in a single sequencing run. In more specific terms, each cluster on the flow cell produces a single sequencing read. For example, 10,000 clusters on the flow cell would produce 10,000 single reads and 20,000 paired-end reads.

**Sanger sequencing:** The sequencing method, also known as capillary electrophoresis sequencing, developed in 1977 by Frederick Sanger. It involves a method of DNA sequencing based on the selective incorporation of chain-terminating dideoxynucleotides by DNA polymerase during in vitro DNA replication.

**sensitivity:** In sequencing, the ability to detect low-frequency variants.

**sequencing by synthesis (SBS):** SBS technology uses 4 fluorescently labeled nucleotides to sequence the tens of millions of clusters on the flow cell surface in parallel. During each sequencing cycle, a single labeled dNTP is added to the nucleic acid chain. The nucleotide label serves as a “reversible terminator” for polymerization: after dNTP incorporation, the fluorescent dye is identified through laser excitation and imaging, then enzymatically cleaved to allow the next round of incorporation. As all 4 reversible terminator-bound dNTPs (A, C, T, G) are present, natural competition minimizes incorporation bias. Base calls are made directly from signal intensity measurements during each cycle, which greatly reduces raw error rates compared to other technologies. The result is highly accurate base-by-base sequencing that eliminates sequence-context-specific errors, enabling robust base calling across the genome, including repetitive sequence regions and within homopolymers.

**sequencing panel:** The subset of genes or target regions identified as regions of interest for a specific area of research.

**somatic mosaicism:** Somatic mosaicism occurs when the somatic cells of the body are of more than one genotype. Mosaicism can result from a single fertilized egg cell, due to mitotic errors at first or later cleavages. It can also arise from a mutation during development, which is then propagated to only a subset of the adult cells.

**target region:** A specific sequence of the genome, identified as a region of interest, due to possible involvement in or association with biological development, pathogenesis, or other area of study of interest to the investigator. The sequence can be a gene, a gene segment, a gene fusion, a promotor region, part of an intron or exon, or any stretch of sequence of interest to the investigator.

**targeted resequencing:** A subset of genes or regions of the genome are isolated and selectively enriched or amplified before sequencing. Targeted approaches using next-generation sequencing (NGS) allow researchers to focus time, expenses, and data analysis on specific areas of interest. Such targeted analysis can include the exome (the protein-coding portion of the genome), specific genes of interest (custom content), targets within genes, or mitochondrial DNA.
References


Industry-Leading Solutions
A Community of Support

From library prep, arrays, and sequencing to informatics, Illumina genomic solutions empower researchers and clinical researchers across the globe to find the answers they seek.

When you join the Illumina community, you become part of a dynamic scientific movement that includes thousands of researchers and industry thought leaders. Throughout the year, we host user group meetings, symposia, consortia, online forums, and other initiatives—all designed to bring the best minds together to share ideas and advance science.

In addition to on-site training, ongoing support, and phone consults, we offer webinars and courses at various Illumina locations. We’re here with all the resources you need to accelerate progress.

Contact your Illumina representative to learn more or visit www.illumina.com/targeted-resequencing

A global genomics leader, Illumina provides complete next-generation sequencing workflow solutions to the basic and translational research communities. Illumina technology is responsible for generating more than 90% of the world’s sequencing data.* Through collaborative innovation, Illumina is fueling groundbreaking advancements in the fields of oncology, reproductive health, genetic disease, microbiology, agriculture, and forensic science.